Aspects of the regulation of expression of *pcnB*, which encodes poly(A) polymerase I of *Escherichia coli*

Nigel P. Binns

A thesis presented for the degree of Ph.D.

Institute of Cell and Molecular Biology
University of Edinburgh

December 2000

To Alison and my parents

In the end, you always end up writing what you can, not what you want. The idea in your head is always better.

Shelagh Stephenson, October 1998

Contents

Acknowledgements

First and foremost, I would like to thank Alison, my wife, for sharing her life with me. Also for her patience during the course of this work, especially towards the 'end'. It has been tested to and beyond its tolerance more often than I (or she) may care to remember. I would also like to thank my parents for providing the background that enabled me to get this far. Though neither are scientists themselves, as my Father says, it must be in their genes! A special thank you is reserved for Medhat Khattar. His help, advice encouragement and especially his friendship have kept me going throughout my post-graduate training. I am also indebted to him for introducing me to the delights of Kebab Mahal, Edinburgh's finest curry house.

Several people contributed to the work described in this thesis. My thanks to Uta Binnie for her work on RNA I half-life determination and quantitative Northern analysis, Charles Turnbough for sharing his knowledge of reiterative transcription with me and Richard 'the β-gal kid' Smith for teaching me everything he knows about β-galactosidase assays.

Many thanks also to Millie Masters, my supervisor, for allowing me the opportunity to carry out research in her laboratory, the MRC for funding this work, Willie Donachie for allowing me to bounce some ideas off him and Frank and Graeme for their excellent photographic work.

Abstract

The *pcnB* gene encodes poly(A) polymerase I (PAP I), the major *Escherichia coli* poly(A) polymerase involved in mRNA processing, and the decay of small plasmid copy number control RNAs. Prior to the study presented here, the transcriptional and translational organisation of the *pcnB* region was ill-defined. In this work, the *pcnB* promoter, identified by sub-cloning and primer extension analysis, was found to resemble closely a consensus σ^{70} promoter in both sequence and activity. Operon fusions of *pcnB* with *lacZ* are only slightly derepressed in a Δ*pcnB* background, providing little evidence of autoregulation at the level of transcription. Translation of the *pcnB* message was shown, by site-directed mutagenesis, to commence from the ribonucleotides AUU, a triplet that is very rarely used as an initiation codon. Furthermore, it was demonstrated that at the level of translation, the *in vivo* expression of *pcnB* is specifically subject to negative regulation, over a four-fold range, by Initiation Factor-3 (IF-3), whose own translation is also initiated from an AUU codon. Additional tests showed that a single chromosomal copy of wild-type *infC* (which encodes IF-3) can induce maximal IF-3-mediated repression of *pcnB*. Apart from the Shine-Dalgarno region, *pcnB* appears to lack any of the known sequence determinants encoded by *infC* that are believed to facilitate the use of AUU as an initiation codon. Several *pcnB* homologues are identified from other eubacteria that potentially could utilise an AUU triplet as a translational initiation codon. A conserved run of T residues (five to seven nucleotides long) exactly nine nucleotides downstream from the -10 region of these homologues suggested that the expression of *pcnB* might be subject to regulation by pyrimidine-sensitive selection of transcriptional start sites, coupled with UTP-dependent reiterative transcription. This possibility was investigated.

Abbreviations

ADP	Adenosine-5'-diphosphate
Amp	Ampicillin
ATP	Adenosine-5'-triphosphate
bp	Base pair(s)
cAMP	3',5'-cyclic adenosine monophosphate
cfu	Colony forming units
Cmp	Chloramphenicol
CTP	Cytidine-5'-triphosphate
(d)dATP	2'(3'-di)-deoxyadenosine-5'-triphosphate
(d)dCTP	2'(3'-di)-deoxycytidine-5'-triphosphate
(d)dGTP	2'(3'-di)-deoxyguanosine-5'-triphosphate
(d)dTTP	2'(3'-di)-deoxythymidine-5'-triphosphate
(d)dNTP	2'(3'-di)-deoxyribonucleoside-5'-triphosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dsDNA	Double-stranded deoxyribonucleic acid
EDTA	Diaminoethanetetra-acetic acid
GTP	Guanosine-5'-triphosphate
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid)
IPTG	Isopropyl-β-D-thiogalactoside
Kan	Kanamycin
kb	Kilobase pair(s)
kDa	Kilodalton
moi	Multiplicity of infection
MOPS	Morpholinopropanesulphonic acid
mRNA	Messenger ribonucleic acid
OD	Optical density
ONPG	o-nitrophenyl-β-D-galactoside
PEG	Polyethylene glycol
PIPES	Piperazine-N,N'bis (2-ethanesulphonic acid)
R	Resistant
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute

S	Sensitive
SDS	Sodium dodecyl sulphate
ssDNA	Single-stranded deoxyribonucleic acid
TEMED	N,N,N',N'-tetramethyl-1,2-diaminoethane
Tet	Tetracycline
Triton X-100	Octylphenoxypolyethoxyethanol
Tris	2-amino-2-hydroxymethyl-1,3-propandiol
ts	Temperature-sensitive
TTP	Thymidine-5'-triphosphate
UTP	Uridine-5'-triphosphate
UV	Ultraviolet
v/v	Volume by volume
w/v	Weight by volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

CHAPTER 1

Introduction

1 Introduction

The work described in this thesis is principally concerned with two questions. These are the organisation of the *pcnB* locus in *Escherichia coli* and the way in which the synthesis of its product, poly(A) polymerase I, is controlled. PcnB was discovered because of its role in the control of plasmid ColE1 replication, where it proved to affect the rate at which a small anti-sense inhibitor of replication (RNA I) decays. PcnB was later shown to have a role in mRNA decay as well. Translation of *pcnB* proved to be initiated at a rare codon and because of this, the synthesis of its gene product is under an unusual sort of control, involving initiation factor 3 (IF-3) as a regulatory element. This introduction will be divided into three sections, covering each of the topics mentioned above: ColE1 replication control and organisation of the *pcnB* locus, PcnB and mRNA stability and the autoregulation of IF-3.

1.1 ColE1 copy number control and the polyadenylating activity of PcnB

The regulation of plasmid replication depends on a number of host and plasmid encoded factors. For most of the plasmids that are found in Gram negative bacteria, copy number control is exerted through one of two principal mechanisms of negative regulation. Some plasmids (examples include ColE1 and R1) encode a small diffusible RNA molecule which blocks a critical step in replication by functioning as an antisense transcript. Alternatively, some plasmids (such as F, P1 and pSC101) employ multiple binding sites, in the form of direct repeats known as *iterons*, that compete with the origin of replication for a plasmid encoded initiator protein (Chattoraj, 2000). It is this fundamental difference in replication control that appears to determine whether plasmids from otherwise unrelated groups are responsive to the activity of the host-encoded plasmid copy number control gene product, PcnB. This protein alters the stability of the small RNA molecule that controls the frequency of plasmid replication and hence the copy number. In a *pcnB* strain, the stability of this RNA is increased, causing a reduction in the frequency of replication.

1.1.1 Plasmid maintenance

The replication of plasmids has been studied extensively as it was believed that they offered a simple model for understanding the process of bacterial chromosome replication. Because of the diverse range and nature of plasmids, this belief is no longer tenable. However, the study of plasmid replication control mechanisms has evolved into a discipline in its own right and is important for a variety of reasons, not least because plasmids are the primary agents of horizontal gene transfer in bacteria. As such they are responsible not only for the scourge of infectious drug resistance, but are essential for the technology of genetic engineering! (Cohen, 1993).

Plasmids are duplex, supercoiled DNA molecules which, in *Escherichia coli*, range in size from about 5 kb to 100 kb (equivalent to approximately 0.1-2% of the *E. coli* chromosome). They are stable genetic elements that exist in an extrachromosomal state, and are able to replicate in an autonomous and self-controlled way (Sherratt, 1974). Such control is always exerted at the replication initiation stage, presumably because the initiation events, as opposed to the elongation and termination steps, are invariably replicon specific (del Solar *et al*, 1998).

For a given bacterial host under specific growth conditions, a particular plasmid is maintained at a defined intracellular concentration, or copy number. This characteristic can be used to broadly divide plasmids into two groups; those that are maintained at only a few copies per cell and those that are maintained at a much higher copy number. Members of the former group tend to be large and because of their low copy number, they require a partition system to ensure accurate segregation at the time of cell division. Members of the latter group are generally small and do not encode an active partition system. Chance inheritance of the plasmid by a daughter cell after division is facilitated by the high copy number of the plasmid (Helinski *et al*, 1996).

The maintenance of plasmid copy number is determined by the frequency of initiation of plasmid replication such that, on average, one replication event occurs per plasmid copy per generation. A variety of mechanisms peculiar to individual plasmid species are employed to ensure their maintenance in a host cell at a characteristic copy number. These control mechanisms are

responsible for the phenomenon of incompatibility (Tomizawa and Itoh, 1981), which was the original basis of plasmid classification; two phenotypically distinguishable plasmids that share the same replication control mechanisms are designated incompatible and are not stably co-inherited. The random selection of a high copy number plasmid for replication, and the subsequent random assortment for partitioning are the two features of plasmid maintenance responsible for this unstable co-inheritance (Helinski *et al*, 1996).

Initiation of plasmid replication usually requires at least one self-encoded protein that specifically recognises the plasmid's own origin of replication, although this is not universally true. However, all plasmids are dependent upon host-cell proteins for the replication of their DNA and synthesis of the proteins they encode, and so are unable to propagate outside the cell.

1.1.2 ColE1 replication

The replication and maintenance of copy number in ColE1 and related plasmids (such as the naturally occurring pMB1 and the recombinant derivative, pBR322) have been investigated in considerable detail (reviewed in del Solar and Espinosa, 2000; Wagner and Brantl, 1998; del Solar *et al*, 1998; Helinski *et al*, 1996; Wagner and Simons, 1994; Cesareni *et al.*, 1991 and references therein). This small (6 kb) double-stranded circular DNA molecule is maintained at a high copy number averaging 15 per cell. ColE1 replication proceeds in a uni-directional mode via a θ mechanism (Wickner, 1978). Replication is initiated from a unique origin contained within a minimal 500 bp replicon. Initiation does not involve permanent nicking of either strand (Blair, and Helinski, 1975). Unlike most plasmids, ColE1 does not encode a specific replication initiation protein, and hence the process is entirely dependent upon host-encoded proteins. *In vitro* studies suggest a minimal requirement for RNA polymerase, RNase H and DNA polymerase I in ColE1 replication.

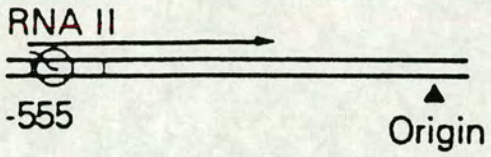
Up to three proteins bind near the origin and bring about supercoiling of the circular molecule (Blair and Helinski, 1975), thereby exposing single-stranded regions to which *E. coli* RNA polymerase binds (Backman *et al*, 1978). The replication pre-primer (RNA II) is transcribed from a promoter 555 nucleotides upstream from the origin, *oriV* (Figure 1.1.1).

Figure 1.1.1 ColE1 primer formation

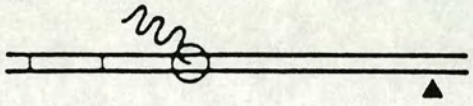
Straight lines represent double-strands of DNA; wavy lines, RNA transcripts; small circles, RNA polymerase. The DNA template forms an eye structure from the origin of replication (filled triangle). One alternative of step 3 causes primer hybridisation, and the other (shown in parentheses) does not. Arrows in the eye structures in steps 5 and 6 indicate cleavage of hybridised RNA by RNase H. The thick lines in the eye structures are newly synthesised DNA strands. From Eguchi *et al* (1991).

FORMATION OF PRIMER AND ITS REMOVAL

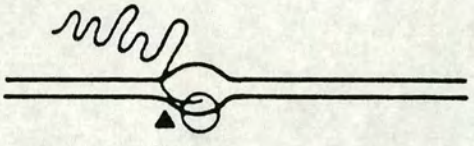
1, Transcription initiation by RNA polymerase



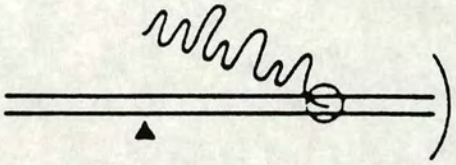
2, Elongation of RNA II



3, RNA/DNA hybrid formation, coupling

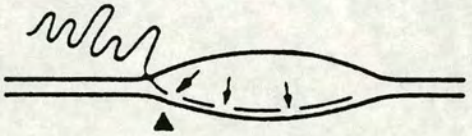


(Absence of coupling


)

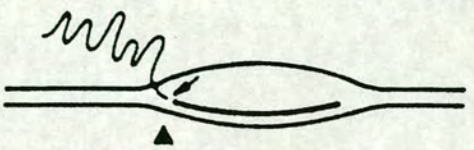
4, Elongation of hybrid

5, Cleavage by RNase H



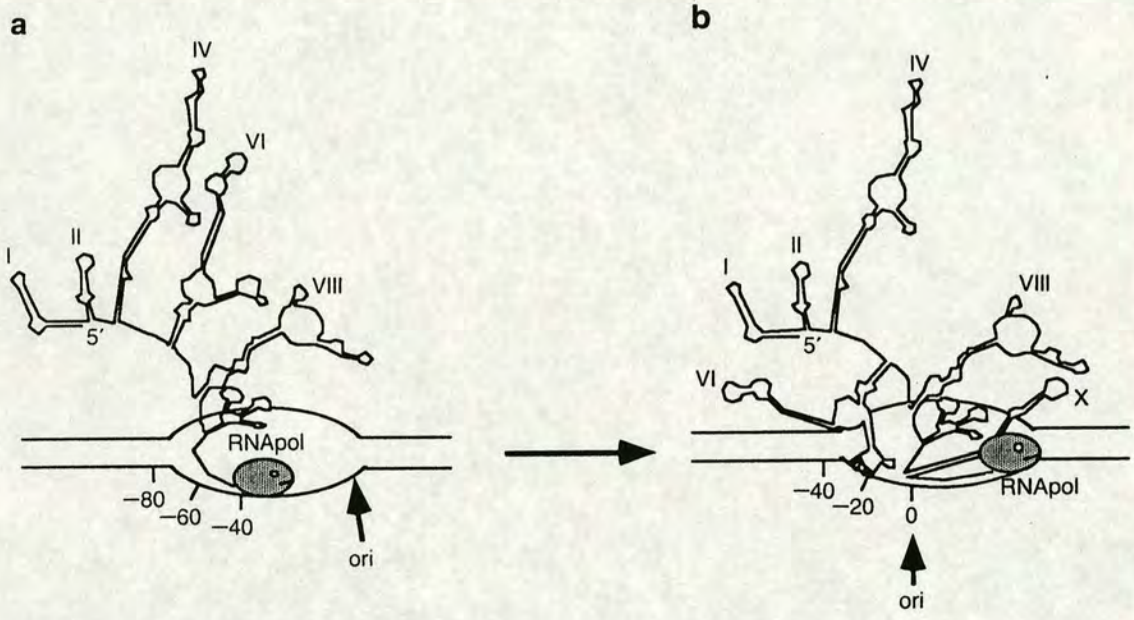
6, Addition of dNMP by DNA polymerase I

7, Elongation of DNA & removal of primer



This transcript, whose 5' end adopts a complicated 'αβ' structure (Figure 1.1.4) with at least 10 stems and loops, is unusual in that it resists release from its template at the origin. It was shown by mutational analysis that the RNA II/DNA template hybrid persists due to a specific hydrogen bonding interaction between the pre-primer and its template (a G-rich loop region in RNA II at -265 and a C-rich region on the template just upstream from the origin at -20). The regions involved are brought into close proximity as a consequence of the secondary structure of RNA II (Figure 1.1.2).

Figure 1.1.2 Persistent RNA II-DNA template hybrid formation.



Persistent hybrid formation between RNA II and the DNA template at the origin. (a) Secondary structure of the RNA II after RNA polymerase has transcribed approximately 500 nucleotides. (b) Hydrogen bonding between an RNA II loop at position -260 and the DNA template at position -20. From Cesareni *et al* 1991).

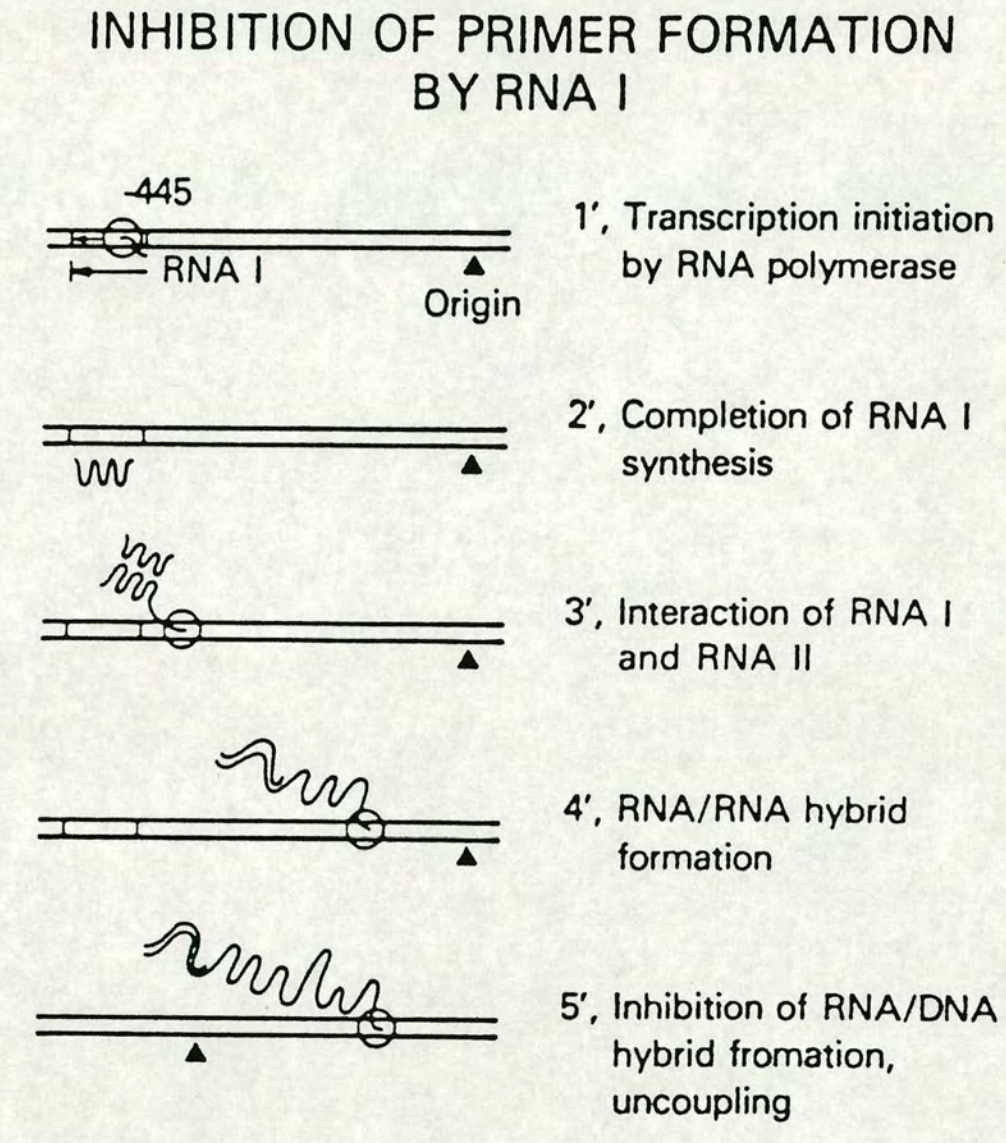
Due to the RNA II/DNA hybrid, a single strand displacement D-loop is formed (Itoh and Tomizawa, 1978; Masukata and Tomizawa, 1984; Marians, 1992). The hybrid is sensitive to RNase H, which fragments the nascent RNA producing several primers that are extended by *E. coli* DNA polymerase I, with the help of DnaB helicase. This leads to the synthesis of a 6S piece of DNA (400 nucleotides) joined to a 20 ribonucleotide long primer, the whole remaining associated in the D-loop structure (Itoh and Tomizawa, 1978;

Marians, 1992). The DNA polymerase III replicating complex takes over leading strand synthesis from DNA polymerase I to extend the fork (Itoh and Tomizawa, 1978; Blair and Helinski, 1975). Lagging strand synthesis can be initiated from a PriA binding site on the light strand (Masai and Arai, 1988).

1.1.3 Control of ColE1 copy number: synthesis and interaction of RNA I and RNA II

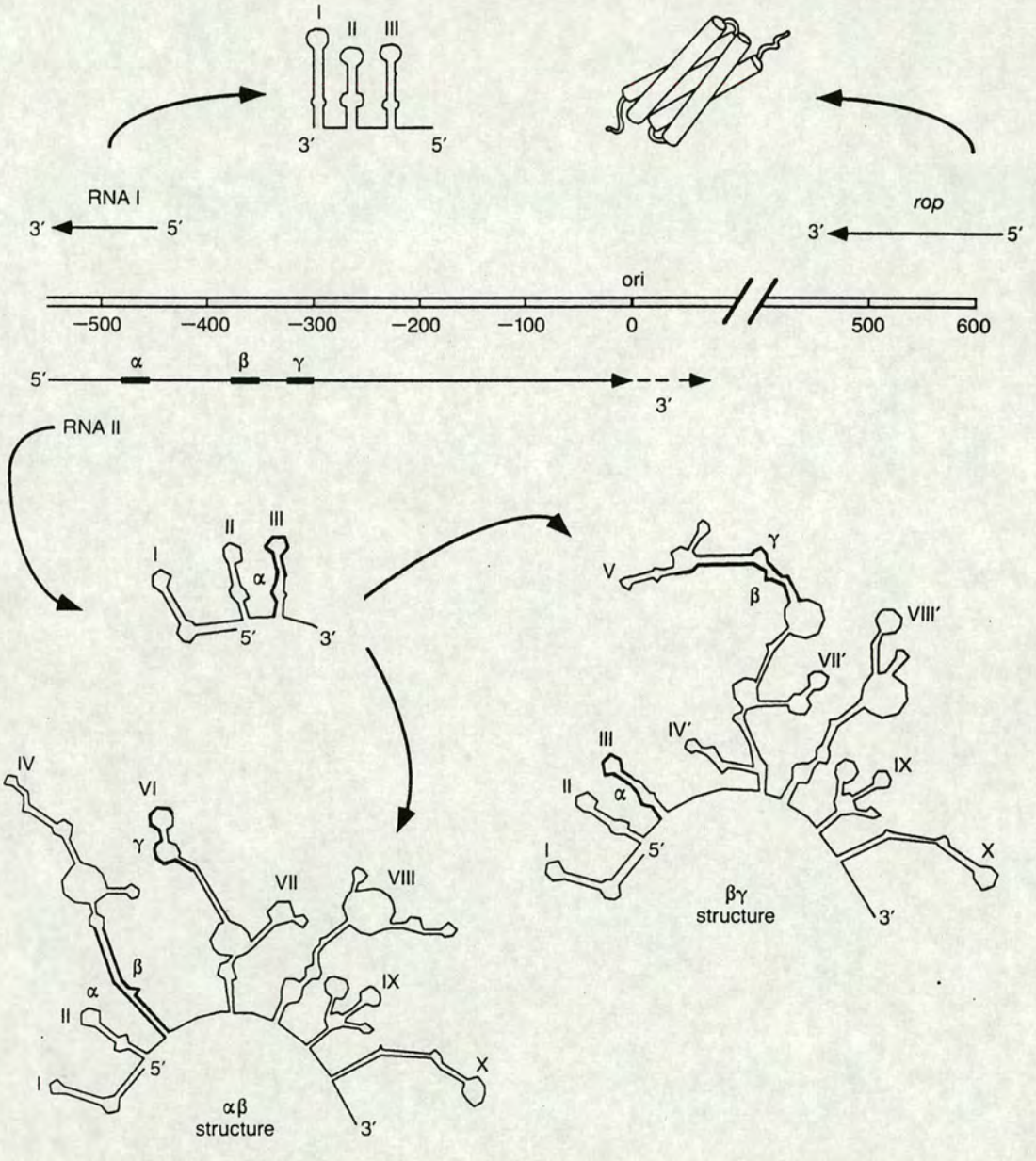
The regulation of ColE1 copy number (as for other replicons) is mediated at the stage of replication initiation by a plasmid-specific system. Initiation is dependent upon the ability of RNA II to adopt the αβ conformation that allows it to form the stable hybrid with its template. The process is largely independent of the rate of pre-primer synthesis (Brenner and Tomizawa, 1991), although the presence of three GATC sites in the RNA II promoter may allow the control of primer transcription via methylation (Polisky, 1988). Initiation of replication is known to be negatively regulated by two molecular species (reviewed in Eguchi *et al*, 1991; Simons and Kleckner, 1988 and references therein). The principal regulator is a 108 nucleotide antisense molecule called RNA I, which is also a substrate for PcnB. This RNA is transcribed from nucleotides -455 to -555 on the opposite strand to RNA II and is therefore complementary to the 5′ end of the RNA II pre-primer (Figure 1.1.3). RNA I is a freely diffusable *trans*-acting factor that is capable of interacting with any RNA II as it is transcribed. This interaction causes RNA II to adopt an alternative 'βγ' conformation that is incapable of forming the stable primer/template hybrid with *oriV*, necessary for replication initiation (Figure 1.1.4).

Figure 1.1.3 Inhibition of ColE1 primer formation by RNA I



Straight lines represent double-strands of DNA; wavy lines, RNA transcripts; small circles, RNA polymerase; filled triangle, origin of replication. Binding of RNA I to RNA II in step 3' inhibits formation of the RNA/DNA hybrid and consequently primer formation. From Eguchi *et al* (1991).

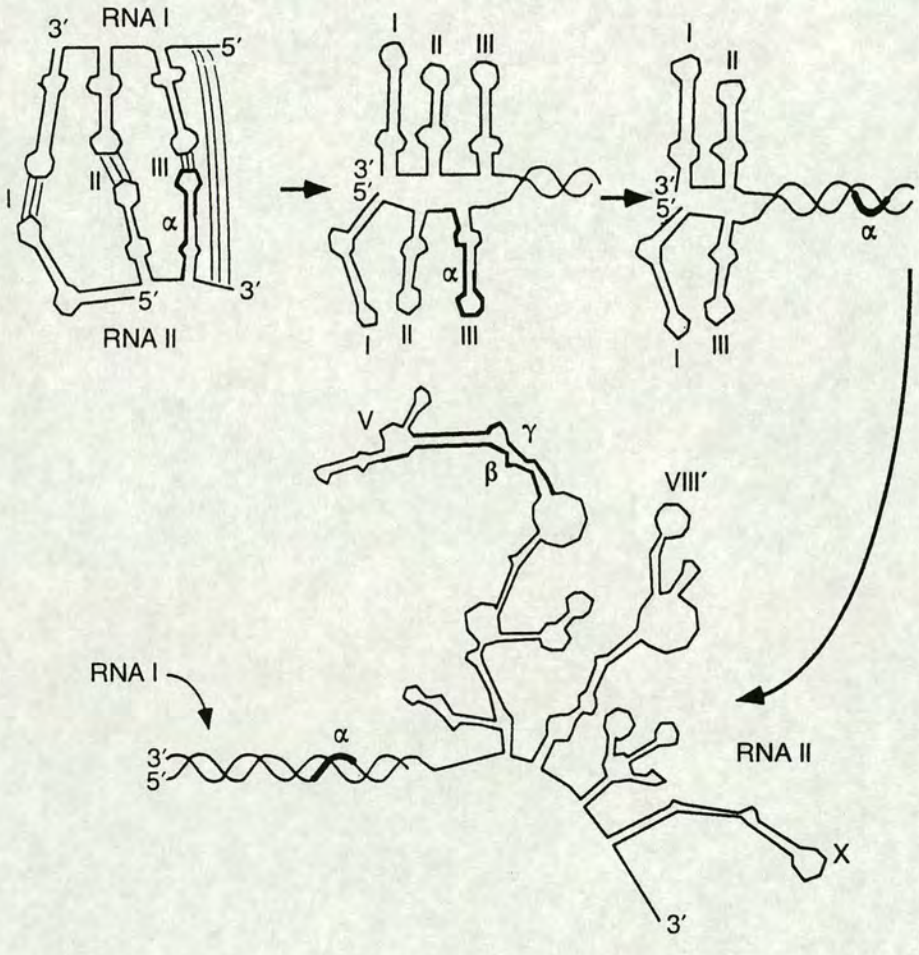
Figure 1.1.4 The αβ and βγ conformations of RNA II



The αβ structure and the βγ structure are the two functionally important alternative conformations adopted by RNA II that influence its ability to form the persistent RNA-DNA hybrid at the origin. The genetic map of the region essential for initiation of ColE1 DNA replication and its control is also shown, along with the structure of the two inhibitors of plasmid replication, RNA I and Rom. From Cesareni *et al* (1991).

The so called 'kissing' and subsequent pairing interaction between RNA I and RNA II takes place between the three corresponding stem and loop structures adopted by both species (Wagner and Brantl, 1998). RNA I, which itself assumes a complex secondary structure, interacts with loops I-IV at the 5′ end of RNA II, preventing the formation of loops VI-X and thereby interferes with primer formation (Figure 1.1.5).

Figure 1.1.5 The stepwise process of binding of ColE1 RNA I to RNA II



RNA I and RNA II interact at the loops of their folded structures. This interaction facilitates pairing that starts at the 5′-end of RNA I. The pairing propagates progressively as the stem-loop structures unfold, and finally, RNA I hybridises along the entire length to RNA II. The interaction between RNA I and RNA II results in the formation of a hybrid molecule that sequesters the α sequence in RNA II and prevents the formation of the αβ conformation leading to inhibition of primer formation. From Cesareni *et al* (1991).

RNA I is an effective regulator of plasmid copy number because the amount present in the cell is directly determined by the number of plasmid origins. The constitutive expression, unstable nature, and rapid turnover of RNA I (it has a half-life of about two minutes) means that it is able to respond quickly to changes in plasmid copy number during steady-state growth. RNA I is maintained at a constant concentration that is in excess of RNA II. (approximately 150-fold; 1000 nM to 7 nM; Brenner and Tomizawa, 1991). This ensures that the concentration of RNA I is essentially unaltered by its interaction with RNA II. The constant concentration of RNA I is maintained by a feedback mechanism that inversely changes the replication frequency (and hence the plasmid copy number) in response to changes in the concentration of RNA I. It is probable that this negative control of ColE1 initiation can be explained by the inhibitor dilution model of Pritchard *et al*, 1969 (though factors other than inhibitor concentration are thought to regulate the maximum rate of replication; Thomas, 1988). The inhibitor (RNA I) is synthesised by the replicon. Its concentration (which determines the initiation rate) is therefore proportional to the plasmid concentration. As the cell volume increases during growth, so the inhibitor's concentration decreases by dilution, and the initiation block is removed allowing the replication of new plasmid molecules. Apart from its dilution as a result of the growth-dependent increase in cell volume, the amount of RNA I present in the cell fluctuates due to enzymatic degradation and the loss and synthesis of new plasmid molecules after cell division.

The second negative regulator of plasmid replication is a 63 amino acid plasmid-encoded protein known as Rom (for **R**NA **o**ne **m**odulator), which maps immediately downstream from the origin (Cesareni *et al*, 1982). Plasmids in which the *rom* gene has been deleted exhibit around a five-fold elevation in copy number (Twigg and Sherratt, 1980). This *trans*-acting protein appears to stabilise the RNA I/RNA II complex by reorganising the interaction between the two species (Tomizawa, 1990). In so doing, it may increase the rate of complex formation. This is critical to the negative regulatory activity of RNA I, because once RNA II has been extended to longer than 360 nucleotides, loops VI-X can no longer be disrupted and the primer becomes refractory to the activity of RNA I. Rom appears to bind to the 3′ end or mid region of RNA II thereby extending the period of interaction between RNA I and RNA II. In addition, Rom also helps align

RNA I and RNA II (Masai and Arai, 1988), but appears to play no role in plasmid incompatiblity. A single point mutation introduced within the replication primer of ColE1 has been shown to mimic the high copy number phenotype of the pUC class of vectors (Lin-Chao *et al*, 1992). This Rom suppressible lesion is believed to destabilise RNA I/RNA II complex formation in a temperature-sensitive manner.

Unlike other plasmids (and the *E. coli* chromosome), ColE1 can undergo initiation of replication in the absence of *de novo* protein synthesis. This fact has been exploited for the purpose of amplifying (that is, increasing the plasmid/chromosome ratio) ColE1-derived cloning vectors such as pBR322, by growing *E. coli* in the presence of chloramphenicol. The copy number of such 'relaxed' plasmids can be increased from about 20 to 3000 copies per chromosome per bacterium by the cessation of protein synthesis (Clewell, 1972). It is also possible that a reduction in the intracellular concentration of Rom contributes to the elevation in copy number. 'Stringent' plasmids such as pSC101, that are dependent on protein synthesis for their replication, cannot be amplified by this method.

Control of R1 copy number

In addition to affecting the copy number of ColE1-related plasmids, PcnB is also known to modulate the replication frequency of plasmid R1, which belongs to the unrelated IncFII group. As with ColE1, PcnB alters the turnover rate of an unstable antisense RNA that controls the frequency of R1 replication (Söderbom *et al*, 1997).

1.1.4 Controlling ColE1 copy number: decay of RNA I and copy number control

1.1.4.1 RNA I stability and Ribonuclease E

As the half-life of RNA I is likely to be a significant factor in determining the frequency of ColE1 replication initiation, it was anticipated that ribonuclease activity would have a major role to play in this process. RNase E, a 114 kDa endoribonuclease encoded or regulated by the *rne/ams* locus (**a**ltered **m**RNA **s**tability), has been implicated in bulk mRNA stability (Mudd *et al*, 1990). At the non-permissive temperature, RNA I isolated from *rne* ts mutants has a half-life of 16 minutes compared with two minutes for a wild type isogenic strain (Lin-Chao and Cohen, 1991); a decrease in plasmid replication shadows the elevated half-life indicating that a mutation in the *rne* locus affects both the chemical stability and biological activity of RNA I. RNase E cleaves RNA I near its 5′ end. The major cleavage product, pRNA I_{-5}, retains the ability to bind to RNA II. Earlier work reported that the five nucleotides cleaved from the 5′ end of RNA I were essential for its negative regulatory activity (Tomizawa, 1984; Tamm and Polisky, 1985). However a mutationally created homologue of the RNase E cleavage product (pppRNA I_{-5}) was able to prevent the priming of ColE1 replication both by wild-type RNA II and by its own complement, RNA II_{-5} (Lin-Chao and Cohen, 1991). Under slow growth conditions, an observed 10-fold elevation in the intracellular concentration of pppRNA I_{-5} was linked to a proportionate decrease in plasmid copy number. This observation suggests that copy number is modulated by the rapid degradation of pRNA I_{-5}, rather than by the loss *per se*, of the five terminal nucleotides removed by RNase E. It was subsequently shown that RNA I's susceptibility to degradation by RNase E requires the presence of an unpaired 5′ nucleotide tail (Bouvet and Belasco, 1992); the addition of a synthetic stem-loop to the 5′ end diminishes this susceptibility. The 5′ mono-phosphorylated form of the RNA I cleavage product decays at a faster rate than its tri-phosphorylated homologue, suggesting that the phosphorylation state of the terminal 5′ nucleotide of RNA I may be an important determinant of its half-life (Lin-Chao and Cohen, 1991).

In the last decade it has been discovered that RNase E does not normally work in isolation but is, rather, part of a complex termed the *degradosome* (Carpousis *et al*, 1994) which includes, in addition to RNase E, a 3' to 5' exonuclease, polynucleotide phosphorylase (PNPase, encoded by *pnp*), a helicase and other components. Whereas PNPase degrades the RNA I product of RNase E cleavage, pRNA I_{-5}, efficiently, it degrades pppRNA I_{-5} much more slowly (Xu and Cohen, 1995), suggesting that the 5' end structure of RNA as well as the 3' structure is important in the functioning of this nuclease.

1.1.4.2 PcnB: a positive regulator of ColE1 copy number

Not until 1986 was a chromosomal locus identified that affected the rate of initiation of replication of ColE1, although a mutation in the *polA* polymerase domain completely prevents plasmid replication. Whilst working on an unrelated subject, a group isolated a class of host-determined ColE1-related copy number mutants (Lopilato *et al*, 1986). They named this locus, which mapped to 3.4 minutes on the *E. coli* chromosome, *pcnB* (for **P**lasmid **C**opy **N**umber). Subsequently, two other groups independently identified this gene (Liu and Parkinson, 1989; March *et al*, 1989); its nucleotide sequence, first determined by Liu and Parkinson (1989), was later corrected (Blattner *et al*, 1997). Hosts mutated at this locus were typically found to exhibit an up to 10-fold reduction in plasmid copy number (by single-cell ampicillin resistance and direct measurement of plasmid DNA content) for this recessive phenotype. A *pcnB* deletion mutant grows normally on rich media, suggesting that the protein has no unique essential cellular role (Masters *et al*, 1993).

From the inferred amino acid sequence, the product of the *pcnB* gene was deduced to be a 48 kDa (410 residue) globular protein; this molecular weight is consistent with its rate of migration on SDS-PAGE sizing gels. It has a high basic amino acid composition characteristic of proteins that interact directly with nucleic acids (Liu and Parkinson, 1989). Localised sequence homology with *E. coli* tRNA nucleotidyltransferase initially suggested that PcnB might function by interfering directly with RNA I/RNA II complex formation, as RNA I adopts a secondary structure similar to that of tRNAs (Masters *et al*, 1990), but this was shown not to be the case (He *et al*, 1993).

The biochemical function of PcnB was identified serendipitously. A group investigating the biosynthesis and function of bacterial polyadenylated mRNA isolated a poly(A) polymerase from *E. coli* (Cao and Sarkar, 1992). The gene was located using degenerate oligonucleotides based on its N-terminal sequence. This revealed an exact match with a region upstream of the previously hypothesized *pcnB* open reading frame, suggesting that the correct reading frame begins further upstream than first thought. Experiments to be described in this thesis confirm this interpretation.

The N-terminal sequence of PcnB determined by Cao and Sarkar (1992) begins with a lysine residue. However, the codon that determines this amino acid is never used for translational initiation. According to these workers, the nearest possible start site, UUG, is 17 amino acids upstream of the lysine codon, which suggests either that artefactual *in vitro* cleavage or *in vivo* post-translational modification of the polypeptide may have occurred. The putative UUG translation start site is infrequently used as an initiation codon in *E. coli*. The utilisation of UUG as an intiation codon has been implicated as a factor that can reduce the translational efficiency of a gene (Reddy *et al*, 1985).

The rapid degradation of pRNA I_{-5} seen in *pcnB*$^+$ strains is not observed in *pcnB* cells; the half-lives of pRNA I_{-5} and pppRNA I_{-5} are elevated 10-fold in the latter (Xu *et al*, 1993). The accumulation of biologically active RNA I decay products in a *pcnB* strain indicates that PcnB is required for their swift breakdown (Xu *et al*, 1993; He *et al*, 1993). The identification of PcnB as Poly(A) polymerase I, combined with the observed extended half-life of RNA I in *pcnB* mutants suggests that PcnB normally adenylates RNA I and that this adenylation is necessary for its rapid decay. Although it has proved difficult to demonstrate the existence of adenylated RNA I in wild-type cells, adenylation of RNA I has been demonstrated in cells in which PcnB is expressed at higher than normal levels, or in which RNase III is absent (Binnie *et al*, 1999). Lengthened RNA I, presumed adenylated, can also be visualized in cells deficient in PNPase.

How might adenylation of RNA I expedite its decay? To understand this we need to consider the mode of action of PNPase, the 3' to 5' exonuclease implicated in RNA I decay. [A deficiency of RNase II, the second exonuclease involved in mRNA decay, does not reduce RNA I decay (Binnie *et al*, 1999)].

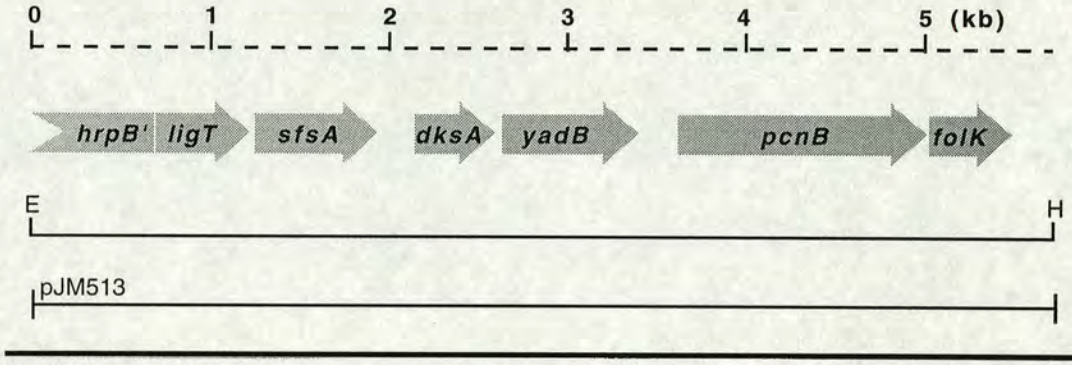
This will be done in more detail in the next section, but for the purposes of this section it is sufficient to point out that unpaired nucleotides (i.e. a single-stranded region) at the 3' -end of a potential substrate appear to be necessary for these exonucleases to bind RNA. RNA I, as synthesized, has a strong stem-loop at its 3' -end, with only two nucleotides beyond it. These are apparently insufficient for efficient binding and potentiation of PNPase (or RNase II) action.

Interest in ribonucleases has rapidly expanded with the realisation that RNA degradation is as important as transcription in determining the levels of gene expression. Ribonuclease activity has roles in RNA metabolism, processing, turnover and degradation (Deutscher, 1988; Petersen, 1992). The bulk of currently acquired knowledge has been derived from studies on *E. coli* because of the relative ease with which appropriate mutants can be isolated (Deutscher, 1988). A consistent problem with this approach has been the lack of an observable mutant phenotype (Deutscher, 1993a). For this reason, current strategies involve construction of double, triple and sometimes even quadruple mutants at known exoribonuclease loci (Deutscher, 1993b). The unifying observation from such studies has been the overlapping specificities of many of these enzymes (Kelly and Deutscher, 1992). Although each appears to have a preferred substrate, many are capable of compensating (though not always as efficiently) for the absence of another.

1.1.5 Genetic organisation of the *pcnB* locus

Most of the genetic loci located around the 3.4 minute region of the *E. coli* chromosome have been functionally defined to some extent (Berlyn, 1998; Blattner *et al*, 1997; Fujita *et al*, 1994; Riley, 1993). A 5.7 kb fragment cloned from the region (creating pJM513) spans five complete open reading frames, including *pcnB*, that have been biochemically and/or genetically characterised and one, *yadB*, who's function remains unknown (Figure 1.1.1). Because genes with related functions are often found clustered together on the *E. coli* chromosome - either as single or polycistronic transcription units - further insight into the functioning of PcnB might be gained by examining its flanking regions.

Figure 1.1.1 Genetic organisation around the 3.4 minute region of the *E. coli* chromosome.



Genetic organisation around the 3.4 minute region of the *E. coli* chromosome. pJM513 is a pBR328 derivative that contains the indicated 5.7 kb insert. E, *Eco*R I; H, *Hind* III. Map is drawn to scale.

Upstream and distal to *pcnB* is the *ligT* locus. This open reading frame encodes a 20 kDa protein that exhibits a 2'-5' RNA ligase activity (Arn and Abelson, 1996). This enzyme specifically ligates tRNA half-molecules (that result from the endonucleolytic removal of introns from eukaryotic tRNA pre-cursors) containing nucleoside base modification and shows a preference among different tRNA species. It acts in the absence of ATP to form a 2'-5' phosphodiester linkage. As no introns of any kind are found in the full genomic complement of tRNA genes in *E. coli*, its role in this organism remains unclear.

The 3' end of *ligT* abuts the start of the *sfsA* open reading frame. Its gene product is a 26 kDa protein believed to be a regulatory factor involved in maltose metabolism (Kawamukai *et al*, 1991). Overexpression of this gene increases amylomaltase levels.

Immediately downstream from *pcnB* is the *folK* locus. This gene encodes 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase, an 18 kDa enzyme that catalyses the first step in dihydrofolate biosynthesis (Talarico *et al*, 1992; Talarico *et al*, 1991). As its presumptive translation start codon overlaps the stop codon for *pcnB* (cgcATGAcag), it is likely that these two genes are co-transcribed. The failure of *pcnB* deletants missing this region to grow well at high temperature on minimal medium, a phenotype suppressed by

expression of *folK in trans*, supports this hypothesis (S. McAteer, *pers. comm.*). However, as with the other genes discussed in this section, there is no evidence to suggest that FolK is concerned with the maintenance of plasmid copy number. There remains the possibility, however, that the activity of FolK is metabolically related to the poly(A) polymerase activity of PcnB. This idea is expanded in Chapter 5.

The *dksA* locus lies in the region between *sfsA* and *pcnB*. This non-essential gene encodes a 17.4 kDa protein that was identified as a dosage-dependent suppressor of a *dnaK* deletion mutant (Kang and Craig, 1990). It suppresses not only the temperature-sensitive growth but also the filamentous phenotype of the *dnaK* deletion strain; defective growth of lambda phage is not suppressed.

An open reading frame of unknown function designated *yadB* is located in the *dksA-pcnB* intergenic region (Masters *et al*, 1992). Its hypothetical gene product is a 33.6 kDa protein that shows similarity over 25 kDa to the N-terminal half of glutamyl tRNA synthetase (encoded by the gene *gltX*; Breton *et al*, 1986). It is unclear whether *yadB* is an active cistron. The possibility that it has the capacity to be expressed is examined in this thesis and the findings are discussed in Chapter 5.

1.2 The polyadenylating activity of PAP I and its role in RNA decay

1.2.1 Eubacterial mRNA degradation

The bacterial way of life involves rapid response to environmental change. Such rapid response requires that proteins be made quickly when needed but not made when not required. This is achieved in part by rapidly synthesising mRNA in response to changing conditions, so as to allow the requisite protein synthesis, and by destroying it immediately after use. This rapid mRNA turnover is mediated by a group of dedicated enzymes, principally ribonucleases, the individual roles of which have come under close scrutiny in the last decade. (Deutscher, 1993; Petersen, 1992; Deutscher, 1988). With the discovery in the mid 1990s of the degradosome, a molecular machine active in mRNA decay, interest in RNA decay has been further stimulated. Much of the material in this section has been distilled from the following recent reviews (Steege, 2000; Carpousis *et al*, 1999, Coburn and Mackie, 1999).

The first step in the synthesis of a protein is the synthesis of its corresponding mRNA. The need for gene-specific mRNA synthesis provides an opportunity for controlling the level at which a gene product will be made. Control of transcription and of translation have long been recognized and studied as key points at which control of gene expression can be exerted. More recently it has come to be realised that, while bulk mRNA turns over rapidly, individual mRNAs can have differing half-lives. Interest in the mechanism of mRNA decay has also increased with the realisation that the rate of mRNA degradation can be important in determining levels of gene expression.

Ribonucleases fall into one of two categories. Endonucleases cleave within RNA molecules; exonucleases digest in a processive fashion from RNA ends. In bacteria mRNA degradation appears to be initiated by one or more endonucleolytic cleavages, followed by digestion with 3' to 5' exonucleases. No 5' to 3' exonucleases, which do exist in eukaryotes, have been identified in any bacterial species. *E. coli* has over 20 RNases. Most of these, however, are involved in rRNA and tRNA synthesis and processing with only six thought to be involved in mRNA decay. The table below lists these enzymes.

Table 1.2.1 mRNA decay enzymes

Enzyme	Gene	Subunit size (kDa)	Enzyme activity	Reference
		Endoribonucleases		
Ribonuclease E (RNase E)	*rne*	118	**endo**nuclease (single strand)	Coburn *et al*, 1999
Ribonuclease III (RNase III)	*rnc*	25	**endo**nuclease (double strand)	Dunn, 1976
RNase G	*rng* (*cafA*)	55	overlaps that of RNaseE	Tock *et al*, 2000
		3' to 5' Exoribonucleases		
Ribonuclease II (RNase II)	*rnb*	72.5	Hydrolytic; removes NMPs	Gupta *et al*,1977
Polynucleotide phosphorylase (PNPase)	*pnp*	77	Phosphorolytic; removes NDPs	Portier, 1975
RNase R	*rnr* (*vacB*)	92	Hydrolytic	Cheng *et al*, 1998
Oligoribonuclease	*orn*	20.7	2-5 nt to single nt	Ghosh and Deutscher, 1999

1.2.1.1 The endonucleases

The major endonuclease of RNA decay is RNase E, a large, autoregulated protein with multiple domains. Only the N-terminal 60% of the protein, which has the RNase activity, appears to be required for cell survival (Kido *et al*, 1996). The N-terminal portion of RNase E contains an RNA binding fold first described for ribosomal protein S1; this fold is also present in PNPase and RNase II (Bycroft *et al*, 1997). RNase E cleaves 5' to AU in AU rich sites, but does not appear to show strict sequence specificity. It is likely that it recognises some structural characteristic which is not yet completely defined. Cleavage occurs in single-stranded regions only. A 5' end of at least three unpaired nucleotides (Bouvet and Belasco, 1992), preferably mono- rather than triphosphorylated (Lin-Chao and Cohen, 1991), is also required for

RNase E activity. The C-terminal end of RNase E enables it to interact with other nucleases, in particular with PNPase. This will be considered further below, in the discussion of the degradosome.

The recently identified RNase G is paralogous to RNase E and, like it, has a role in processing ribosomal RNA; a specific role in mRNA decay has not yet been demonstrated. It, too, is essential for viability.

A third RNA endonuclease, RNase III, which again has a role in rRNA processing, cuts RNAs in double stranded regions; the structural motif recognized is not yet clear. Specific instances of mRNA substrates have been described (see Kushner, 1996 for refs.); interestingly RNase III is involved in the degradation of the mRNA specifying the exonuclease PNPase. RNase III deletion mutants remain viable, but are much reduced in growth rate.

1.2.1.2 The exonucleases

After cleavage by endonuclease, RNA degradation proceeds via 3' to 5' cutting by exonucleaseses, principally polynucleotide phosphorylase (PNPase) and RNase II. PNPase is a phosphorolytic enzyme; in the presence of inorganic phosphate it catalyses the processive 3' to 5' phosphorolytic degradation of RNA to yield nucleotide 5' diphosphates. The enzyme is a trimer composed of 711 amino acid "alpha" subunits (Portier, 1975). Depending on the purification method, it may copurify with what were initially regarded as "beta" enzyme subunits, but which have since been shown to be dispensable for enzyme activity and identical to the enzyme enolase. The fact that enolase is part of the degradosome perhaps explains this association.

RNase II, a hydrolytic enzyme that removes nucleoside 5' monophosphates from the 3' end of RNA, accounts for 90% of the exonucleolytic activity on mRNA in *E. coli* (Deutscher & Reuven, 1991) (with the remainder attributable to PNPase). Curiously, since it accounts for so great a fraction of RNase activity, inactivation of *rnb* has little phenotypic effect other than an increase in the stability of certain RNAs. The probable explanation of this observation is that RNase II which, like PNPase, is impeded by secondary structures such as stem-loops, will remove residues until it encounters inhibiting secondary structure and then disassociate from the RNA. This leaves a modified

molecule which, since it has no single-stranded tail at all, is then inaccessible to PNPase as well as to RNase II (Coburn and Mackie, 1996a).

The *E. coli* genome encodes a paralog of RNase II, RNase R. This enzyme appears to be responsible for residual hydrolytic activity in *rnb* mutants. Either RNase II or PNPase can be inactivated and cells will still grow quite well. However, although a deficiency of either of these enzymes does not compromise viability, inactivation of both appears to be lethal (Donovan and Kushner, 1986). Curiously, since RNase R accounts for much less of the hydrolytic activity than does RNase II, inactivation of *rnr* cannot be achieved in *pnp* cells either. Since both RNase II and RNase R can be inactivated together, this suggests that each of these enzymes has a unique function which can be carried out by PNPase, but not by the other hydrolytic exonuclease! (Cheng *et al*, 1998). The relationships amongst the roles of these three enzymes is clearly far from fully understood.

The final RNase currently believed to have a role in RNA decay is oligonuclease This enzyme breaks down the two to five nucleotide long RNA molecules which are the products of exonuclease action. This enzyme appears to be essential (Ghosh and Deutscher, 1999).

1.2.1.3 The degradosome

The degradosome is a multi-protein complex which degrades RNA. It was discovered during attempts to purify the elusive RNase E. It was found that PNPase co-purifies with RNaseE, and that additional proteins were copurified as well (Carpousis *et al*, 1994). During an attempt to identify proteins which bind to a stem-loop RNA sequence, REP, a binding activity with a molecular weight in excess of 500 kDa was identified (Py *et al*, 1994). This proved to be a complex containing PNPase, RNase E and other proteins, suggesting the existence of a dedicated RNA degrading structure which might bring together endo- and exonucleolytic activities. Amino-terminal sequencing of the unknown proteins in these complexes eventually identified enolase, the ATP dependent RNA helicase, RhlB, and, in lesser amounts, polyphosphate kinase. It was subsequently discovered that, under conditions likely to activate the helicase, RNA degrading activity past refractory stem-loop structures was expedited (Py *et al*, 1996; Coburn *et al*,

1999). Functions for the enolase and polyphosphate kinase components of the degradosome have not yet emerged.

Characterisation of protein interactions in the degradosome has shown that the C-terminal portion of RNase E probably acts as a scaffold for organizing the interactions of the other proteins within the degradosome. RNase E has separate and distinct binding sites for PNPase, enolase and RhlB within its C-terminal half (Vanzo *et al*, 1998) and distinct from its nucleolytic activity, which resides in its N-terminal portion. There are still many unresolved questions related to degradosome function. These include the stoichiometry and interaction of the component proteins, the fraction of the component proteins which are found in the structure (as opposed to free in the cytoplasm), and whether there is a specific subcellular localization for the complex. Another question needing an answer is for what fraction of RNA degradation it is responsible. Answers to these questions should be forthcoming in the next few years.

1.2.2 PAP I and the role of polyadenylation in RNA decay

A poly(A) polymerase activity was described in *E. coli* as early as 1962 (August *et al*, 1962), well before eukaryotic messages were shown to have poly(A) tails. Indeed, *E. coli* long served as the commercial source of this enzyme. However, with the discovery that eukaryotic mRNAs carry poly(A) tails up to 200 nt long, while few bacterial messages were polyadenylated at all (and, if they were, tail lengths were very short) it became generally accepted that polyadenylation in bacteria lacked functional significance. The identification of PcnB as PAP I resulted in a resurgence of interest in the role of bacterial polyadenylation (Cao and Sarkar, 1992). The demonstration that RNA I decay was much reduced in the absence of PAP I led to an examination of the role of polyadenylation in mRNA decay in several different systems (Hajnsdorf *et al*, 1994a; O'Hara *et al*, 1995). It appears that the addition of a poly(A) tail can stimulate the rate of degradation by PNPase or by RNase II by about five-fold *in vitro* (Coburn and Mackie, 1996b). The unstructured 3' end that polyadenylation provides is thought to promote the binding of these exonucleases to otherwise refractory substrates (Littauer and Soreq, 1982; Coburn and Mackie, 1996a; Coburn and Mackie, 1996b).

Coburn and Mackie, in their comprehensive 1999 review of mRNA decay in *E. coli*, present several models of mRNA decay in which PAP I plays a role. In their *5' tethering model* it is supposed that the degradosome binds at the 5' end of the target RNA and also at a second site within the molecule, cutting at the latter site and, while remaining bound to the 3' end of the cut, degrades the RNA between it and the initial 5' site of binding. This process would be repeated with decay occurring in an overall 5' to 3' direction.

Degradation of the 3' end of the molecule, especially when this end has a stem-loop which would render it refractory to exonuclease action, could be facilitated by a poly(A) extension acting to "feed" the 3' OH end to PNPase within the degradosome. That this is not a model sufficient for all situations is made clear by the observation that poly(A) extensions are found not only at the 3' OH ends of transcripts, but at internal sites as well, suggesting that fragments generated during decay are adenylated in their turn.

The alternative *3' latent decay* and *3' tethering models*, invented to consider the special problems of transcripts with 5' stem loops or other structural features which would inhibit 5' binding of the degradosome, incorporate this observation. These models hypothesize that the initial step in decay is adenylation by PAP I. In the 3' latent decay model, the participation of the degradosome is not necessary. Adenylation is followed by PNPase degradation: a poly(A) tail of sufficient length allows the conversion of PNPase from the non-processive to the processive form, in which it can take advantage of RNA "breathing" (intermittent relaxing of base-pairing at the base of stem-loops) to invade and degrade RNA in H-bonded stems and so to degrade from 3' to 5'. Any dissociation of PNPase from its substrate could provide further opportunity for polyadenylation at internal sites.

In the 3' tethering model initial polyadenylation enables the degradosome to associate with the substrate at its 3' end via PNPase binding. During subsequent degradation by PNPase, the 5' end of the target becomes associated with the RNase E portion of the structure and degradation then proceeds as in the 5' tethering model.

These models, although providing an interesting take off point for future research, do not incorporate all known facts. Foremost amongst these is the fact that *pcnB* is not an essential gene. Indeed, deletion mutants of this gene

show no growth disadvantage at all (Masters *et al*, 1993). The preliminary identification of a second poly(A) polymerase (PAP II; Cao *et al*, 1996; Kalapos *et al*, 1994), which might share PAP function, was apparently premature; there is no candidate PAP II known at present (Mohanty and Kushner, 1999).

Thus, PAP I is an enzyme useful in RNA decay, but, in so far as enzyme catalyzed RNA decay is an essential process, polyadenylation is not an essential part of it.

1.3 Expression and negative autoregulation of IF-3

1.3.1 Translation initiation

The work described in this thesis presents evidence that the expression of *pcnB* is subject to regulation by IF-3 at the level of translation. For this reason, a short review of translational initiation, with particular reference to the role of IF-3, is presented here.

In addition to the ribosome and mRNA, translation initiation in bacteria, (reviewed in Kozak, 1999; McCarthy and Brimacombe, 1994; McCarthy and Gualerzi, 1990; Jacques and Dreyfus, 1990; Gualerzi and Pon, 1990; Gold, 1988), requires a specific tRNA ($tRNA_f^{Met}$), at least three initiation factors, and the hydrolysis of a GTP molecule. The three characterised initiation factors , IF-1, IF-2 and IF-3, bind to the 30S ribosomal subunit. IF-1 and IF-2 binding is dependent on the binding of IF-3. They all (IF-1 only in combination with IF-2 and IF-3) stimulate the rate of ternary complex (30S-mRNA-$tRNA_f^{Met}$) formation. IF-3 also inhibits association between free 30S and 50S ribosomal subunits and promotes dissociation of ternary complexes with noninitator tRNAs. IF-2 is a GTP-binding protein that recognises, and probably positions, f-Met-$tRNA_f^{Met}$ in the ternary complex. A kinetic scheme, similar to that applied to transcription, has been proposed for translation initiation (Gualerzi and Pon, 1990). The 30S subunit, to which initiation factors are fixed, binds the ligands (mRNA and fMet-$tRNA_f^{Me}$) in either order to form a ternary preinitiation complex with the ligands not interacting in a productive way. A rate-limiting first-order conformational change permits the formation of the active initiation complex. Upon binding of the 50S subunit, IF-3 and IF-1 leave the active initiation complex and the 70S initiation complex forms. Finally, IF-2 promotes positioning of the fMet-$tRNA_f^{Met}$ in the P site and is ejected from the complex. GTP is hydrolyzed and translation elongation starts.

Translation initiates on a specific region of the mRNA that is characterised by the presence of an initiator codon located five to nine nucleotides downstream from the Shine-Dalgarno sequence. Together these two elements, plus the bases in between, define the minimal ribosome binding site (RBS). From its complete genome sequence, translation in *E. coli* is predicted to initiate at AUG, GUG or UUG codons for 83, 14 and 3% of its

genes, respectively (Blattner *et al*, 1997). The three to nine nucleotide Shine-Dalgarno sequence has been shown to base pair with the 3' end of 16S rRNA (a component of the 30S ribosomal subunit) during initiation (Hui and deBoer, 1987). Sequences in messenger RNA other than the SD sequence are suspected of being capable of pairing with 16S rRNA (Sprengart and Porter, 1997; Sprengart *et al*, 1990; Thanaraj & Pandit, 1989). Statistical analysis of translation initiation sites shows a nonrandom distribution of nucleotides between the position -20 to +14. This corresponds almost exactly to the region of mRNA protected from ribonuclease digestion by bound ribosomes (the RBS) and to the end of the reverse transcriptase extension from primers annealed to the mRNA downstream from the RBS (Gold, 1988). Sequence flanking the RBS can also influence the efficiency of initiation. The extent of this sequence defines a functional region known as the translational initiation region (TIR). Secondary structure within mRNA plays an important role in the selection of authentic initiation sites and affects the efficiency of translation at these sites. In general, secondary structure hinders initiation at authentic translation start sites. Expression of the coat protein of RNA phage MS2 can be quantitatively correlated to the stability of a hairpin structure containing the RBS (de Smit and van Duin, 1990a). Some of the characteristics of initiation sites are interdependent. For instance, it has been shown that a strong Shine-Dalgarno interaction can compensate for a structured initiation region (de Smit and van Duin, 1994). Conformational constraints imposed by the formation of secondary structure ensures that the initiation of translation from non-authentic start sites does not occur (de Smit and van Duin, 1990b).

1.3.2 Initiation Factor 3

Translation initiation factor IF-3 is required for peptide chain initiation in *E. coli*. It performs two important functions. Firstly, it stabilises free 30S ribosomal subunits, by binding directly to free 30S subunits that are released from the pool of 70S ribosomes. The presence of IF-3 prevents 30S subunits from re-associating with 50S subunits, thus ensuring a constant supply of free 30S subunits for initiation complex formation. In this capacity, it is acting as an anti-association factor. The reaction between IF-3 and the 30S subunit is stoichiometric: one molecule of IF-3 binds per subunit. As there is a relatively small amount of IF-3 - one molecule per seven ribosomes - its availability determines the number of free 30S subunits (Howe and Hershey, 1983).

Secondly, it favours ternary complex formation by participating in the kinetic selection of the correct initiator region of mRNA, and destabilising initiation complexes containing non-initiator tRNAs.

IF-3 is the gene product of *infC*, which in *E. coli* is located at 38 minutes on the chromosome and encodes a 20.4 kDa protein. The gene for IF-3 has been sequenced in several bacterial species, including *E. coli*, and all start with exceedingly rare initiation codons. These are mostly AUU codons, (Sarcedot *et al*, 1982; Pon *et al*, 1989; Liveris *et al*, 1993; Hu *et al*, 1993) but AUC has been found in the *dsg* gene that encodes a protein similar to IF-3 in *M. xanthus* (Cheng *et al*, 1994). The expression of IF-3 is negatively autoregulated at the translational level, and a change of the AUU initiation codon to AUG was shown to abolish this negative auto-control (Butler *et al*, 1986; Butler *et al*, 1987). This clearly shows that the unusual AUU codon is necessary for regulation. Further experiments indicate that if the normal AUG initiation codon of several genes is changed to AUU, these genes become regulated by IF-3 levels in the cell (Sacerdot *et al*, 1996; Sussman *et al*, 1996). This indicates that the AUU initiation codon behaves as a translational operator for the *infC* gene, and is not only necessary but may even be sufficient for regulation. If the classical regulation model by which the repressor binds to its own mRNA and affects ribosome attachment were valid in this case, IF-3 should specifically recognise and bind to an AUU sequence when placed in the context of an authentic TIR. However, regulation by IF-3 levels can be obtained with other abnormal initiation codons besides AUU (Sacerdot *et al*, 1996; Sussman *et al*, 1996). These can vary at any of the three bases of the initiation codon. Thus, it is not AUU *per se* that seems to be necessary and sufficient for regulation by IF-3, but the presence of a non-canonical initiation codon in the TIR. Since IF-3 regulation is found with initiation codons that differ from AUU at every position, there is no specific nucleotide that is responsible for control, and thus it seems unlikely that IF-3 binds to its mRNA. Since IF-3 is known to bind to the 30S subunit, it is reasonable to propose that IF-3 regulates its expression *in vivo* through its interaction with the ribosome rather than directly with its mRNA. Recent work demonstrating that IF-3 is capable of inducing structural changes in the 30S ribosomal subunit (either in the free state or as part of the ternary initiation complex) seems to support this idea (Shapkina *et al*, 2000; de Cock *et al*, 1999).

Once on the ribosome, the factor is able to differentiate between standard and non-standard initiation codons.

Early *in vitro* experiments showed that IF-3 stimulates translation initiation with the usual initiation codon, AUG, and inhibits initiation with some other codons (Berkhout *et al*, 1986), in particular AUU (Lateana *et al*, 1993). Experiments using the toeprint technique have shown that IF-3 selectively favours the binding of initiator tRNA versus elongator tRNA at the RBS (Hartz *et al*, 1990). This selectivity is only found with AUG, GUG, and UUG initiation codons but not with AUU. These experiments show quite clearly that IF-3 is able to 'inspect' the initiation codon and the anticodon domain of the initiator tRNA. This faculty of IF-3 to discriminate against initiation on non-canonical codons could be due to either direct recognition of the two last bases of the codon and their cognate bases on the anticodon, or to some ability to detect codon-anticodon complementarity. Recent work suggests that this discrimination is dependent upon the ability of IF-3 to recognise simple mismatches in the codon-anticodon interaction at the second and third base of the codon (Meinnel, *et al* 1999). It thus seems that, in this case, regulation is explained by the activity of the factor. When the 30S subunit attaches to an mRNA with a canonical initiation codon, IF-3 somehow accelerates the rate at which the preternary complex changes to a productive ternary complex. If the 30S subunit attaches to the *infC* mRNA with its AUU initiation codon, IF-3 instead accelerates dissociation of the preternary complex, thus down regulating its translation. *In vitro* experiments have shown that IF-2 dependence is increased when translation is initiated with AUU instead of AUG (Lateana *et al*, 1993). The same set of experiments indicate that IF-3-stimulated initiation at AUG and IF-3-stimulated dissociation at AUU codons give superimposable curves that saturate at an IF-3/30S ratio of one. Both the stimulation and the dissociation are reduced with rRNA mutants known to be defective in IF-3 binding. This indicates that stimulation of AUG-dependent initiation and dissociation of initiation complexes formed with AUU are due to IF-3 binding to the same site on the ribosome.

1.3.3 Mechanism of *infC* negative autoregulation

The molecular nature of the mechanism that determines the negative autoregulation of *infC* is inherently linked to the mechanism that determines the coordinate expression of ribosomes and IF-3, and is dependent on the unusual AUU initiation codon in *infC* (Brombach and Pon, 1987; Butler *et al*, 1986).

The *in vivo* level of IF-3 has been shown to vary proportionately with cellular growth rate, resulting in an essentially constant IF-3/ribosome ratio of approximately 1:7 (Howe and Hershey, 1983). The expression of ribosomes and initiation factors is coordinately regulated such that the ratios are kept constant at a variety of growth rates (Howe and Hershey, 1983).

The number of ribosomes present in exponentially growing cells is roughly proportional to the growth rate (Gausing, 1980; Maaloe, 1979). The macromolecular nature of the ribosome requires that the gene products of the ribosomal protein and rRNA genes be coordinately expressed. This is accomplished primarily by the regulation of rRNA gene transcription and autogenous feedback regulation of ribosomal protein mRNA translation (Keener and Nomura, 1996). The precise mechanism for the regulation of rRNA gene transcription is unclear, but it has been suggested that the cellular level of non-translating ribosomes is the feedback signal that determines growth rate control of *rrn* promoters (Keener and Nomura, 1996; Cole *et al*, 1987).

Gold *et al* proposed a model for the autogenous translational regulation of IF-3 expression that is based on the unusual sequence of *infC* mRNA around the rare AUU initiation codon (Gold, 1988; Gold *et al*, 1984). Essentially, the model predicts that high levels of cellular IF-3 would prevent the translation of *infC* mRNA, because IF-3 itself causes the dissociation of mRNA-30S subunit complexes formed with mRNA containing non-canonical initiation codons, such as AUU. At low IF-3 concentrations, an elevated level of IF-3 deficient ribosomes permits the translation of *infC* mRNA. This model suggests a simple, passive mechanism for the coordinate regulation of IF-3 and ribosome levels. As growth rate increases, ribosome content is elevated by virtue of growth rate control of *rrn* gene transcription. This elevation leads to a transitory decrease in the IF-3/ribosome ratio and a concomitant

elevation in the translation of *infC* mRNA by IF3-deficient ribosomes. Ultimately, the IF-3/ribosome ratio is reestablished (at approximately 0.15) and the translation of *infC* mRNA is repressed. Thus, the key cellular parameter in the growth rate regulation of IF-3 synthesis is the IF-3/ribosome ratio. The model predicts that any change that eliminates the autoregulation of IF-3 production, such as a mutation in the initiation codon from AUU to AUG, will also uncouple the expression of IF-3 from growth rate regulation. Because the start codon of *pcnB* is AUU, the possibility that translation initiation of *pcnB* is negatively regulated by IF-3 was examined in this work and the findings are discussed in Chapter 4.

CHAPTER 2

Materials and methods

2.1 Bacterial strains, phage and plasmids

2.1.1 Bacterial strains

Bacterial strains used in this study are listed in Table 2.1.1. These were routinely maintained at 4°C on plates of LB agar or minimal agar plus appropriate supplements. Long-term stocks were preserved in frozen storage buffer at -70°C.

Table 2.1.1 Bacterial strains (*E. coli* K-12, unless otherwise stated)

Strain	Genotype	Source/Reference
BL21 (DE3)	F⁻, *ompT* [*lon*], *hsdS*$_B$ ($r_B^- m_B^-$), *gal*, *dcm*; pLysS (Cpm^R), DE3 = λ (*cI*^{857} *indI*, *sam7*, *nin5 lacUV5*-T7 gene *I*); an *E.coli* B strain	Laboratory stocks
C600	F⁻, *thi1*, *leuB6*, *thr1*, *lacY1*, *hsdR*, *tonA21*, *supE44*	Young and Davis (1983) Promega
CJ236	F', *cat*(=pCJ105; $M13^S$, Cmp^R)/*dut-1*, *ung-1*, *thi-1*, *relA1*, *spoT1*, *mcrA*	Invitrogen
DH5α	F⁻, φ80Δ(*lacZ*)*M15*, Δ(*lacZYA-argF*)*U169*, *deoR*, *recA1*, *endA1*, *hsdR17*, *supE44*, *thi-1*, *gyrA96*, *relA1*, λ⁻	Hanahan (1983) Clontech
IR8903	NM621, Δ*pcnB*::Kan^R	Masters *et al* (1993)
JK378	P90C, *infC362*, *zdh925*::Tn*10* (Tet^R)	Sussman *et al* (1996)
JK382	P90C, *infC*⁺, *zdh925*::Tn*10* (Tet^R)	Sussman *et al* (1996)
JM110	F', *traD36*, *lacI*^{q}, Δ(*lacZ*)*M15*, *proAB*⁺/*rpsL* (Str^R), *thr*, *leu*, *thi*, *lacY*, *galK*, *galT*, *ara*, *fhuA*, *dam*, *dcm*, *supE44*, Δ(*lac-proAB*)	Hanahan (1983) Invitrogen
MDO	*ara-38*(am), Δ (*codB-lac*)*3*, *tsx-68*, *galK42*(am), *galE54*, *trp-49*(am), *relA1*, *spoT1*, λ⁻	*E. coli* Genetic Stock Centre
MDOK	MDO, Δ*pcnB*::Kan^R (from a P1 lysate made on strain IR8903)	This work
MG1655	F⁻, λ⁻, ? fnr⁻	*E. coli* Genetic Stock Centre (1981)

Table 2.1.1 Bacterial strains (continued)

Strain	Genotype	Source/Reference
MG1655K	MG1655, Δ*pcnB*::Kan^R	This work
MM38	F$^-$, *argG6, hisG1, leuB6, metB1, pyrE, gal-6, lacY1, xyl7, supE44, bgl$^+$ fhuA2*, T1^R, Ph80^R, *gyrA* (Nal^R), *rpsL104* (Str^R), *tsx-1* (T6^R), *uhp*, λ-	Masters *et al* (1993)
MM38K	MM38, Δ*pcnB*::Kan^R	Masters *et al* (1993)
MM38KT	MM38, Δ*pcnB*::Kan^R, *zad*::Tn*10* (Tet^R); (from a P1 lysate made on strain MM44)	This work
MM44	Mri84, Δ*pcnB*::Kan^R, *zad*::Tn*10* (Tet^R)	Laboratory stocks
MM378	MM38, *infC362*, *zdh925*::Tn*10* (Tet^R); (from a P1 lysate made on strain JK378)	This work
MM382	MM38, *infC$^+$*, *zdh925*::Tn*10* (Tet^R); (from a P1 lysate made on strain JK382)	This work
Mri84	F$^-$, *ΔlacU169, araD139, thiA, rpsL, relA,* Δ*rbs-7, zad*::Tn*10 pcnB$^+$*	Lopilato *et al* (1986)
NM621	F$^-$, *hsdR, mcrA, mcrB, supE44, recD1009*	Laboratory stocks
OV2	F$^-$, *ara-38*(am), *leu-87*, *lacZ125*(am), *tsx-465*(am), *galK42*(am), *galE54*, *tyrT181*(ts,Sup^{am}), t*rp-49*(am), *his-224, thyA756*, r*elA1*, *spoT1*, *ilv-693*, *deo-84*, λ$^-$	*E. coli* Genetic Stock Centre
P90C	*ara* Δ(lac-pro) *thi*	Miller (1972)
PAP40	MDO, *pcn-40*, *zad*::Tn*10* (Tet^R)	This work
TG1	F' *traD36*, *lacI^q* Δ(*lacZ*)*M15*, *proAB$^+$/supE*, Δ(*hsdM-mcrB*)*5*, *thi*, Δ(*lac-proAB*)	Gibson (1984)

2.1.2 Bacteriophage strains

Bacteriophages used in this study are listed in Table 2.1.2. Phage lysates were stored at 4°C as broth suspensions to which a few drops of chloroform had been added to prevent microbial growth. M13 phage lysates were stored at 4°C but without chloroform, M13 being chloroform sensitive.

Table 2.1.2 Bacteriophages

Bacteriophage	Description	Source/Reference
λ wild-type	Wild-type immunity	Laboratory stocks
λ*vir*	Virulent	Laboratory stocks
λNB21	λRS45, recombined *in vivo* with pNB21, Kan^R	This work
λNB25	λRS45, recombined *in vivo* with pNB25, Kan^R	This work
λRS45	*bla'-lacZ*$_{sc}$, *imm*^{21}, *ninR5*, *ind*$^+$, transcriptional fusion vector with *lacZ* as reporter gene	Simons *et al* (1987)
λRS551	λRS45, recombined *in vivo* with pRS551, Kan^R	This work
λRS552	λRS45, recombined *in vivo* with pRS552, Kan^R	This work
λRS544	mini-*infC'-'lacZ* fusion with the native *infC* AUU start codon altered to CUG	Sussman *et al* (1996)
λRWS945	λRS45, *dnaA'-lacZ* transcriptional fusion contained on a 945 bp *Eco*R1 fragment spanning the *dnaA* promoter region. *lacZ* translated from own start site as stop codons are present in all reading frames between *dnaA'* and *lacZ*	Smith (1995)
M13BM21	M13-based cloning vector with extended polylinker, 7.3 kb	Boehringer Mannheim
M13NB1	1.5 kb *Mlu* I frag. of pNB1 into *Mlu* I site of M13BM21, 8.8 kb	This work

Table 2.1.2 Bacteriophages (continued)

Bacteriophage	Description	Source/Reference
M13NB1C(-2)	M13NB1, C(-2) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1F(2)	M13NB1, F(2) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1I(1)	M13NB1, I(1) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1I(1)M	M13NB1, I(1) of *pcnB* altered to M(1), 8.8 kb	This work
M13NB1L(-7)	M13NB1, L(-7) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1R(4)	M13NB1, R(4) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1R(10)	M13NB1, R(10) of *pcnB* altered to amber, 8.8 kb	This work
M13NB1T(-1)	M13NB1, T(-1) of *pcnB* altered to amber, 8.8 kb	This work
M13NB4	1.3 kb *Bst*E II frag. of pJM516 into *Hin*c II site of M13BM21, 8.6 kb	This work
P1	Wild-type transducing phage	Laboratory stocks
P1 *vir*	Virulent transducing phage	Laboratory stocks

2.1.3 Plasmids

Plasmids used and constructed in the course of this study are listed in Table 2.1.3.

Table 2.1.3 Plasmids

Plasmid	Description	Source/Reference
pACYC184	Tet^R, Cmp^R, p15A replicon, 4.2 kb	Chang and Cohen (1978)
pASHOK	Amp^R, derivative of pKK232-8 (Amersham Pharmacia) with first eight codons from 5' end of *cat* gene deleted and replaced by *lacZ*, 8.1 kb	A. Kumar, *pers. comm.*
pBR322	Amp^R, Tet^R, pMB1 replicon (ColE1-type), 4.4 kb	Bolivar *et al* (1977)
pBR325	Amp^R, Tet^R, Cmp^R, pMB1 replicon (ColE1-type), 6.0 kb	Bolivar (1978)
pBR328	Amp^R, Tet^R, Cmp^R, pMB1 replicon ColE1-type), 4.9 kb	Soberon *et al* (1980)
pHP45Ω	Amp^R, pBR-based with 2.0 kb Ω frag. (Str^R/Spc^R)	Prentki and Krisch (1984)
pJM511	pJM513, deletion derivative with *yadP*, *sfsA*, *dksA*, Amp^R, Tet^S, Cmp^S, 6.8 kb	March *et al* (1989)
pJM513	pBR328, 5.7 kb chromosomal insert with *yadP*, *sfsA*, *dksA*, *yadB*, *pcnB*, *folK*, Amp^R, Tet^S, Cmp^S, 9.4 kb	March *et al* (1989)
pJM516	pJM513, deletion derivative with *pcnB*, *folK*, Amp^R, Tet^S, Cmp^S, 6.8 kb	March *et al* (1989)
pJM516Ω	pJM516, 2.0 kb *Eco*R I Ω frag.of pHP45Ω into *Eco*R I site of pJM516, *pcnB*, *folK*, Amp^R, Str^R, Spc^R, 8.8 kb	This work
pJM513C(-2)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1C(-2) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJM513F(2)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1F(2) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJM513I(1)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1I(1) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work

Table 2.1.3 Plasmids (continued)

Plasmid	Description	Source/Reference
pJM513L(-7)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1L(-7) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJM513R(4)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1R(4) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJM513R(10)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1R(10) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJM513T(-1)	pNB22, 1.5 kb *Mlu* I frag. of M13NB1T(-1) into *Mlu* I site of pNB22, Amp^R, 9.4 kb	This work
pJW30	P_{gal2} from *galE*, transcriptionally fused to *galK*, Amp^R, 3.8 kb	Wright *et al* (1992)
pNB1	pJM513, 419 bp *Kpn* I frag. removed and ends re-ligated, *pcnB*$^-$, Amp^R, 8.9 kb	This work
pNB2	pUC19, 2.7 kb *Cla* I/*Kpn* I frag. of pNB1 into *Acc* I/*Kpn* I sites of pUC19, Amp^R, 5.4 kb	This work
pNB3	pT74, 2.8 kb *Eco*R I/*Hind* III frag. of pNB2 into *Eco*R I/*Hind* III sites of pT7-4, Amp^R, 5.2 kb	This work
pNB4	pASHOK, pMM (a synthetic 39-mer with consensus σ^{70} promoter) into *Bam*HI/*Hin*dIII sites of pASHOK, Amp^R, 8.1 kb	This work
pNB14	pASHOK, 947 bp *Eco*R V frag. of pJM513 into *Sma* I site of pASHOK. *pcnB*' transcriptionally fused to *lacZ*, Amp^R, 9 kb	This work
pNB16	pASHOK, 947 bp *Eco*R V frag. of pJM513 into *Sma* I site of pASHOK. Reverse orientation to pNB14, Amp^R, 9 kb	This work

Table 2.1.3 Plasmids (continued)

Plasmid	Description	Source/Reference
pNB17	pACYC184, 4.9 kb *Bsu*36 I/*Hin*d III frag. of pJM513 into *Bsu*36 I/*Hin*d III sites of pACYC184, Cmp^R, Tet^S, 7.1 kb	This work
pNB18	pASHOK, 0.85 kb *Pst* I/*Hin*d III frag. containing P_{gal2} of pJW30 into *Pst* I/*Hin*d III sites in pASHOK, Amp^R, 8.1 kb	This work
pNB21	pRS551, 2.1 kb *Eco*R I/*Eco*R V frag. of pNB14 into *Eco*R I/*Eco*R V sites of pRS551. *pcnB'* transcriptionally fused to *lacZYA*, Amp^R, Kan^R, 13.3 kb	This work
pNB22	pJM513, 1.5 kb *Mlu* I frag. removed and ends re-ligated, Amp^R, 7.9 kb	This work
pNB23	pNB17, 1.5 kb *Mlu* I frag. removed and ends re-ligated, Cmp^R, Tet^S, 5.5 kb	This work
pNB23C(-2)	pNB23, 1.5 kb *Mlu* I frag. of pJM513C(-2) into *Mlu* I site of pNB23, Cmp^R, Tet^S, 7.1 kb	This work
pNB23F(2)	pNB23, 1.5 kb *Mlu* I frag. of pJM513F(2) into *Mlu* I site of pNB23, Cmp^R, Tet^S, 7.1 kb	This work
pNB23R(4)	pNB23, 1.5 kb *Mlu* I frag. of pJM513R(4) into *Mlu* I site of pNB23, Cmp^R, Tet^S, 7.1 kb	This work
pNB24	pBR328, 2.1 kb *Mfe* I frag. of pJM513 into *Eco*R I site of pBR328, Amp^R, 7kb	This work
pNB25	pRS552, 947 bp *Eco*R V frag. of pJM513 into *Bam*H I site of pRS552. *pcnB'* translationally fused to '*lacZ*, Amp^R, Kan^R, 13.2 kb	This work
pNB26-B	pRS552, 775 bp *Eco*R V/*Mlu* I frag. of M13NB1I(1)M into *Bam*H I site of pRS552. *pcnB'* translationally fused to *'lacZ*, Amp^R, Kan^R, 13.1 kb	This work

Table 2.1.3 Plasmids (continued)

Plasmid	Description	Source/Reference
pNB27	pRS552, 775 bp *Eco*R V/*Mlu* I frag. of M13NB1 into *Bam*H I site of pRS552. *pcnB*' translationally fused to *'lacZ*, Amp^R, Kan^R, 13.1 kb	This work
pRS551	Transcriptional fusion vector containing *lacZ* as reporter gene, pMB1 replicon, Amp^R, Kan^R, 12.5 kb	Simons *et al* (1987)
pRS552	Translational fusion vector containing *lacZ* as reporter gene, pMB1 replicon, Amp^R, Kan^R, 12.3 kb	Simons *et al* (1987)
pSB1	pBR322, *infC+*, AmpR	Butler *et al* (1986)
pT7-4	Amp^R, T7 RNA polymerase/promoter vector, 2.4 kb	Tabor and Richardson (1985)
pUC19	Amp^R, 2.7 kb	Yanisch-Perron *et al* (1985)

2.2 Growth media, buffers and other solutions

Growth media are listed in Table 2.2.1. Antibiotic solutions are listed in Table 2.2.2. Bacterial/phage buffers and other commonly used buffers are listed in Table 2.2.3.

Table 2.2.1 Growth media

LB broth	Difco tryptone	10 g
	Oxoid yeast extract	5 g
	NaCl	10 g
	pH to 7.2 with 1.0 M NaOH	
	Distilled water to 1 litre	
LB agar	LB broth + 15 g Oxoid No. 3 agar per litre	
L.C. top agar	LB broth (made with 5 g NaCl) + 7 g Difco agar per litre	
BBL top agar	BBL Trypticase	10 g
	NaCl	5 g
	Difco agar	6.5 g
	Distilled water to 1 litre	
Nutrient broth (NB)	Oxoid No.2 nutrient broth	25 g
	Distilled water to 1 litre	
	NB has insufficient thymine to support some thymine-requiring strains. It was therefore added as a supplement to a final concentration of 40 $\mu g\ ml^{-1}$	
Nutrient broth agar	Nutrient broth + 15 g Oxoid No. 3 agar	
M9 minimal medium	4x M9 salts	100 ml
	Distilled water	300 ml
	$MgSO_4$ (1.0 M)	0.4 ml
	Carbon Source (20% w/v)	5 ml
M9 salts (4x)	Na_2HPO_4	28 g
	KH_2PO_4	12 g
	NaCl	2 g
	NH_4Cl	4 g
	Distilled water to 1 litre	
M9 minimal agar	4xM9 salts	100 ml
	Oxoid water agar	300 ml
	$MgSO_4$ (1.0 M)	0.4 ml
Oxoid water agar	Oxoid No. 3 agar	20 g
	Distilled water to 1 litre	

Table 2.2.1 Growth media (continued)

MacConkey agar	Peptone	20 g
	Bile salts No.3	1.5 g
	NaCl	5 g
	Neutral red	0.03 g
	Difco agar	15 g
	Distilled water to 1 litre	
VB minimal media	VB salts (20x)	50 ml
	Carbon source (20%)	10 ml
	Thiamine/vitamin B_1 (1 mg ml^{-1})	2 ml
	Supplements as required	
	Distilled water to 1 litre	
VB minimal agar	As VB minimal media + 15 g Difco agar per litre	
20x VB salts	$MgSO_4.7H_2O$	4 g
	Citric acid	40 g
	K_2HPO_4	200 g
	$NaNH_4.HPO_4.4H_2O$	70 g
	Distilled water to 1 litre	

Minimal medium supplements:

A carbon source (glucose unless indicated otherwise) was added to M9 medium to give a final concentration of 0.2%. Stocks were prepared at a concentration of 20% (w/v).

Amino acid supplements were stored in stock solutions of pure amino acids at a concentration of between 2–10 mg ml^{-1} depending upon the solubility of the particular amino acid. Sparingly soluble amino acids, such as tyrosine, were dissolved in 0.01M NaOH. The final concentration of the amino acids in the media was usually in the order of 20–100 μg ml^{-1}.

Pyrimidines were added to minimal media when required. Thymine and uracil were stored at a concentration of 2 mg ml^{-1} in water, and their final concentrations in minimal medium were usually 20–40 μg ml^{-1}.

The only vitamin supplement found necessary in the entire course of this work was thiamine hydrochloride (vitamin B_1). This was stored as a 1 mg ml^{-1} solution in water at 4°C and its final concentration in minimal medium was 2 μg ml^{-1}.

2.2.2 Selection of antibiotic resistance

The routine concentrations for the antibiotics used in this work are shown in Table 2.2.2. All antibiotics were used in both complex and minimal media with the exception of trimethoprim, which was only used in minimal medium.

Table 2.2.2 Antibiotic solutions and other selection agents

Antibiotic	Abbreviation	Solvent	Conc. of stock solution (mg ml^{-1})	Final Conc. in media (μg ml^{-1})
Ampicillin	Amp	Water	100	50-100
Chloramphenicol	Cmp	Ethanol	20	20
Kanamycin sulphate	Kan	Water	25	25-50
Rifampicin	Rif	dimethyl-formamide	100	100
Spectinomycin dihydrochloride	Spc	Water	50	25
Streptomycin sulphate	Str	Water	100	200
Tetracycline hydrochloride	Tet	50 % ethanol	10	10
Trimethoprim	Tmp	Methanol	5	20-50
5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside	X-gal	dimethyl-formamide	20	20-40
Isopropyl-β-D-thiogalactopyranoside	IPTG	Water	20	2

Table 2.2.3 Buffers

Phage buffer	Na_2HPO_4	7 g
	KH_2PO_4	3 g
	NaCl	5 g
	$MgSO_4$ (0.1 M)	10 ml
	$CaCl_2$ (0.1 M)	10 ml
	1% gelatin solution	1 ml
	Distilled water to 1 litre	
Bacterial buffer	KH_2PO_4	3g
	Na_2HPO_4	7 g
	NaCl	4 g
	$MgSO_4$ $7H_2O$	0.2 g
	Distilled water to 1 litre	
TE buffer	*100x solution*:	
	1.0 M Tris-base	
	0.1 M Na_2.EDTA	
	Adjust to pH 8.0 with conc. HCl	
TAE buffer	*50x solution*:	
	2.0M Tris-base	
	1.0 M Na acetate	
	0.05 M Na_2.EDTA	
	Adjust to pH 8.2 with glacial acetic acid	
TBE buffer	*10x solution*:	
	0.89 M Tris-base	
	0.89 M Boric acid	
	0.025 M Na_2.EDTA	
	Adjust to pH 8.2 with 10 M NaoH	
Loading Buffer (DNA)	*5x solution*:	
	Bromophenol blue	40 mg
	Xylene cyanol FF	40 mg
	0.5 M EDTA (pH 8.0)	10 ml
	Ficoll 400	3 g
	Distilled water to 20 ml	

2.3 Bacterial techniques

2.3.1 Growth of bacteria

Bacteria were routinely grown in LB broth or on LB agar at 37°C (or 30°C for temperature-sensitive strains) except where stated. Usually fresh overnight cultures that had been inoculated from a single colony were diluted back the following day and grown as required. Strains harbouring plasmids were grown in medium containing the appropriate antibiotics. VB minimal agar with appropriate carbon sources (at 0.2%), vitamins and amino acids was used for the scoring or supplementation of auxotrophic markers. Medium for the growth of phage λ was supplemented with 10mM $MgSO_4.7H_20$ and 0.2% maltose (w/v) to maximize expression of the λ receptor protein; for phage P1, 2.5 mM $CaCl_2$ was added.

2.3.2 Preparation of competent cells and transformation with plasmid DNA

E. coli cells competent for transformation were prepared using a modification of the method described by Mandel and Higa (1970). 5 ml of LB broth were inoculated with a single colony of a bacterial strain (typically DH5α) and shaken overnight at an appropriate temperature. This culture was diluted 1/100 into fresh LB broth and grown, with good aeration, to an OD_{600} of between 0.3 and 0.4. The culture was chilled on ice, transferred to a universal bottle and centrifuged at 4000 r.p.m. for no more than 5 minutes. The supernatant was discarded and the bacterial pellet resuspended in 1/4 of the original culture volume of an ice cold solution of 0.1M $MgCl_2$. The cells were similarly pelleted again and resuspended this time in 1/25 of the original culture volume of an ice cold solution of 0.1M $CaCl_2$. The cells were kept on ice for at least 30 minutes prior to transformation. At this point the cells could be frozen at –70°C by pelleting and resuspending in 15% (v/v) glycerol in 85 mM $CaCl_2$ or could be used for transformation immediately. Freshly prepared cells always gave the highest transformation efficiency. Plasmid DNA (typically 1–50 ng in < 4 µl) was added to 0.2 ml aliquots of competent cells, mixed and stored on ice for 30 minutes. The cells were heat-shocked for 1.5 minutes at 42°C and returned to ice for 2 minutes. 0.8 ml of LBG (10 ml LB broth + 180 µl 20% (w/v) glucose) was added and the cells incubated at an appropriate temperature for 1 hour to allow expression of plasmid

antibiotic-resistance genes. 0.2 ml of this mixture was then spread onto antibiotic-containing plates and incubated overnight until transformants appeared. Each batch of competent cells was checked by plating a 'no-plasmid' transformation. When transforming with DNA from a ligation reaction, a 'no-insert' control and a 'whole-plasmid' control were transformed in order to monitor the level of self-ligation and the efficiency of the competent cells, respectively.

Using this method, cells were also transformed with replicative form M13 DNA. Several small aliquots (typically in the range 5 to 20 µl) of the heat-shocked phage/cell transformation mix were individually added to tubes containing 0.2 ml of M13-sensitive plating cells. Each tube was mixed with 3 ml of molten BBL top agar, poured onto an LB agar plate and, once set, incubated overnight at 37°C. This procedure should yield discrete plaques of M13 phage.

2.3.3 Transformation by electroporation

For plasmids that were transformed poorly by the $CaCl_2$ method, electroporation was used. 1 litre of LB broth was inoculated with 1/100 volume of a fresh overnight culture of the strain to be transformed. This was shaken at the appropriate temperature until an OD_{600} of 0.5 to 1.0 was reached. The cells were chilled on ice and then washed three times in successively smaller volumes of distilled water (1000 ml, 500 ml and 100 ml) by centrifugation at 4000x g_{max} for 15 minutes at 4°C. The pelleted cells were resuspended in a total of 1 litre distilled water; this was repeated using volumes of 500 and 100 ml distilled water. The cells were then resuspended in 20 ml of 10% (w/w) glycerol in a 30 ml glass Corex tube, centrifuged as above using an SS34 rotor and resuspended in a final volume of 2 ml of 10% (v/v) glycerol. Cells were then either aliquotted and stored at –70°C or used directly for electroporation.

For electroporation, 40 µl aliquots of cells were thawed on ice. DNA was then added (generally 1 to 50 ng in 1 to 2 µl of TE) and the cells were incubated on ice for a further minute. The cells and DNA were then transferred to a pre-chilled electroporation cuvette. Electroporation was performed using a Gene-Pulser™ fitted with a Pulse Controller (Bio Rad Laboratories Ltd.). The 25 µF capacitor was charged to a potential of 2.5 kV and the Pulse Controller set to

200 Ω. The cuvette was pulsed once for a time constant of 4.5–5 ms (field strength 12.5 kV cm^{-1}). The cuvette was then removed from the chamber and the cells were immediately resuspended in 1 ml of SOC (2% Bactotryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20mM glucose). This cell suspension was incubated at an appropriate temperature for 1 hour to allow expression of antibiotic resistance and then dilutions were plated onto the appropriate media.

2.3.4 Single-cell ampicillin resistance assay

The resistance of plasmid-containing strains to ampicillin was used to estimate the relative copy number of the plasmid, since the level of resistance has been reported to be proportional to gene dosage, and therefore reflects plasmid copy number; Uhlin and Nordstrom, 1977). An overnight culture was diluted 1/100 into fresh LB broth plus ampicillin (50 μg ml^{-1}) and grown at 37°C, until an OD_{600} of about 0.3 was achieved. This culture was diluted approximately 10^5-fold in LB broth so as to obtain 200 colonies per antibiotic-free plate, and spread on LB plates containing a range of ampicillin concentrations (typically from 50 μg ml^{-1} to 2000 μg ml^{-1}). After overnight incubation, the colonies were counted and the ampicillin concentration at which 50% of the cells were resistant was determined by plotting the % viability versus the ampicillin concentration.

2.3.5 Frozen storage of bacterial strains

Viable stocks of *E. coli* strains (including those harbouring plasmids) were maintained by long-term storage at -70°C. A fresh 5 ml overnight culture was prepared with antibiotic selection if required. This was centrifuged at 4000 r.p.m. for 10–15 minutes, the supernatant discarded and the cells resuspended in 1 ml of 50% (v/v) glycerol in bacterial buffer. The cells were left on ice for one hour prior to storing at –70°C.

2.4 Phage techniques

2.4.1 Preparation of λ plate lysates

Cells were grown in LB broth containing 20 mM $MgSO_4$ and 0.2% maltose (w/v) at an appropriate temperature until they reached mid-log phase. 0.2 ml aliquots of these cells were then mixed with 10^6 phage, incubated at 37°C for 5 minutes, and 3 ml of molten L.C. top agar containing 20 mM $MgSO_4$ and 0.2% maltose were added. This mixture was poured onto an LB agar plate, left to set and incubated at 37°C overnight or until visible lysis occurred. 5 ml of phage buffer were added to the plate and the layer of top agar scraped off into a sterile 250 ml beaker. A few drops of chloroform were added and the beaker was incubated at room temperature with gentle swirling for 20 minutes. The contents of the beaker were poured into a universal bottle and centrifuged at 4000 r.p.m for 10 minutes in a bench centrifuge. The supernatant was transferred to a fresh bottle and stored over chloroform at 4°C.

2.4.2 Preparation and selection of λ lysogens not conferring antibiotic resistance

A lawn of the bacterial strain to be lysogenised was made by mixing 0.2 ml of a mid-log phase LB broth culture and 3 ml of molten L.C. top agar with 20 mM $MgSO_4$ and 0.2% maltose (w/v). Approximately 200 λ phage particles were added and the mixture poured onto a fresh LB agar plate. Once set, the plate was incubated overnight at 37°C. This should result in the appearance of discrete λ plaques. The centre of a plaque was then touched with a sterile toothpick and streaked out onto another plate, which again was incubated overnight. The resulting single colonies were tested for the presence of the λ phage. A lysogenised bacterium would now be resistant to lysis by phages with the same immunity as the one used to lysogenise the strain, but would be sensitive to λ phages that are virulent, or belong to a different immunity class. This was tested by cross-streaking single colonies of possible lysogens with the original λ phage used for the lysogenisation and λ*vir* as a control. Strains that seemed to be lysogens were finally tested by examining supernatants from LB broth cultures to see if they were producing the correct phage. If the phage λ carried a *lacZ* fusion, lysogens of the phage were

identified by plating out the lysogenised bacterium on a medium containing X-gal.

2.4.3 Preparation and selection of kan^R λ lysogens

Section 3.5.1 describes the Kan^R λ phages, based on those constructed by Simons *et al* (1987), which were used in this work. To isolate recombinants of λRS45 constructed with the kan^r vectors pRS551 and pRS552, 0.1 ml of an overnight culture of an appropriate *lac*$^-$ *E. coli* strain (grown in LB broth plus 20mM $MgSO_4$, 0.2% maltose) was infected with 0.1 ml of various dilutions of the phage lysate to achieve a multiplicity of infection between 10^{-6} and 10^{-2}. The phage were left to adsorb at 30°C without agitation for 20 minutes after which 1 ml of LB broth + 20 mM $MgSO_4$ was added and the mixture was shaken for 2 hours at 37°C. Aliquots of 0.1 ml were spread on LB plates containing kanamycin (25 μg ml^{-1}) and X-gal (40 μg ml^{-1}); the latter provides indication of β-galactosidase synthesis from the phage. Blue kanamycin resistant colonies were selected for further processing from those plates corresponding to the lowest multiplicity of infection.

2.4.4 UV Induction of λ lysogens

Lysogenic bacteria were shaken in LB broth + 20 mM $MgSO_4$ at 37°C until an OD_{540} of 0.3 was achieved. The cells were harvested by centrifugation and resuspended in 7 ml of 20 mM $MgSO_4$. This suspension was transferred to a sterile glass petri dish and the cells were exposed to 600 ergs mm^{-2} sec^{-1} of UV irradiation before being diluted five-fold in fresh LB broth + 20 mM $MgSO_4$. This culture was shaken in the dark at 37°C until lysis occured. This usually took two to three hours. A few drops of chloroform were added, and the lysate clarified by centrifugation prior to titration.

2.4.5 Preparation of phage P1 plate lysates

1ml of late-log phase donor cells grown in 2.5 mM $CaCl_2$ was mixed with 10^6 P1 phage and incubated for 30 minutes at 37°C without shaking. This was added to 3 ml of LB broth with 2.5 mM $CaCl_2$ plus 3 ml L.C. top agar and poured onto a fresh LB agar plate. After overnight incubation, 5 ml of LB broth with 2.5 mM $CaCl_2$ were added to the plate and scraped off with the top agar into a flask. This was shaken with 0.1 ml of chloroform for 20 minutes at 30°C. The lysate was clarified by centrifugation prior to titration. A good lysate would have a titre of up to 10^{11} pfu ml^{-1}.

2.4.6 Phage P1-mediated transduction

The recipient strain of *E. coli* was grown up to late-log phase(OD_{540} ~0.7) in LB broth. The cells were harvested by centrifugation and the bacterial pellet resuspended in 1/10 of the original culture volume of LB broth with 2.5 mM $CaCl_2$. 0.1 ml aliquots of this 10x culture were mixed with 0.1 ml of phage P1 stock and 0.1 ml of a 10-fold dilution of the phage stock. These were incubated for 15 minutes at 37°C. If the desired transductants were to acquire a change in auxotrophy, 0.8 ml of bacterial buffer with 1 mM sodium citrate was added and 0.2 ml aliquots of the cells plated on the appropriate minimal medium. If the desired transductants were to acquire a new antibiotic resistance, then 1 ml of bacterial buffer with 1 mM sodium citrate was added and the cells were centrifuged and resuspended in 1 ml of LBG. This was incubated for 1 hour at 37°C to facilitate expression of the antibiotic resistance gene. 0.2 ml aliquots were plated out on LB agar containing the appropriate antibiotic. Plates were incubated at a suitable temperature until transductants were visible.

2.4.7 Mutagenesis of phage P1 lysates

Random mutations were introduced into the DNA of P1 phage by exposure to hydroxylamine and transferred to the bacterial chromosome by transduction (Hong and Ames, 1971; adapted by Gibbs *et al*, 1992). A high titre P1 lysate (>10^{11} pfu ml^{-1}) was made on a strain encoding a scoreable marker closely linked to the region to be mutagenised. 0.5 ml of this lysate was mixed with 1 ml of phosphate-EDTA buffer (1.0 M K_2HPO_4 added to 1.0 M KH_2PO_4 to give pH 6.0 followed by an equal volume of 10 mM EDTA), 1.5 ml sterile water and 2 ml of freshly prepared hydroxylamine solution (0.35 g NH_2OH plus 560 µl of 4 M NaOH made up to 5 ml with sterile water). After incubation at 37°C for 24 and 40 hours (separate preparations) the phage were precipitated at 4°C in a Sorvall Superspeed centrifuge using an SS34 rotor at 36,000 g for 2.5 hours. The supernatant was removed and the phage particles were recovered from the pellet by adding 0.5 ml of phage buffer and allowing slow resuspension overnight at 4°C.

The viability of a P1 lysate diminishes when it has been treated with hydroxylamine. As such, titration of the mutagenised lysate allowed indirect assessment of the level of mutagenesis by quantifying its effect on plaque formation.

2.5 DNA techniques

2.5.1 Large-scale plasmid preparation

A single colony of the plasmid-carrying bacterial strain was inoculated into 5 ml of LB broth with the appropriate selection and shaken overnight at 37°C . 1 ml of this culture was then used to inoculate 500 ml of LB broth, with similar selection, in a 2 litre flask, which was also shaken overnight at 37°C. The culture was chilled on ice, transferred to two 250 ml centrifuge bottles and centrifuged in a Sorvall Superspeed centrifuge at 5000 r.p.m. for 10 minutes at 4°C. The bacterial pellets were then each washed in 100 ml of TE buffer, pooled to give a total volume of 200 ml, and recentrifuged as above. The resultant cell pellet was resuspended in 5 ml of a 50mM Tris–HCl (pH 8.0), 25% sucrose and transferred to a 50 ml centrifuge tube. 1 ml of lysozyme (20 mg ml^{-1}) was added, the solution mixed, and incubated on ice for 10minutes. 1 ml of 0.5M EDTA(pH 8.0) and 0.8 ml of RNase A solution (10 mg ml^{-1}) were added and incubation continued for a further 10 minutes on ice. Finally 5 ml of a lysis solution containing 100 mM Tris–HCl (pH 8.0), 125 mM EDTA, and 0.2% (w/v) Triton X-100 were added, the solution mixed and incubated on ice for another 10 minutes. The resulting suspension was then centrifuged using a Sorvall SS-34 rotor at 15,000 r.p.m. for 20 minutes at 4°C. The plasmid-containing supernatant could now be subjected to isopycnic gradient ultracentrifugation to separate plasmid and chromosomal DNA.

CsCl (17.1 g) was dissolved in the supernatant. 0.342 ml ethidium bromide solution (10 mg ml^{-1}) was added and the total volume made up to 23 ml with TE. This gave a CsCl density of 1.55 g ml^{-1} and an ethidium bromide concentration of 200 µg ml^{-1}. The solution was then transferred to two 11.5 ml Sorvall Ti-50 crimp-seal centrifuge tubes, balanced to within 0.05 g and then centrifuged in a Sorvall 50-B or 55-B ultracentrifuge at 38,000 r.p.m. for 60 hours at 20°C in a Ti-50 rotor. At the end of the run the tubes were removed from the rotor and the DNA bands visualized using a UV lamp. The lower (denser) plasmid bands were removed from the tubes using a syringe fitted with a wide-bore needle. The two samples were pooled and the ethidium bromide extracted at least five times with isobutanol (isobutanol over CsCl-saturated TE). The sample was dialysed against several changes of

TE (1:2500) at 4°C over a period of 48 hours to remove the CsCl. The plasmid DNA could then be recovered from solution by precipitation.

2.5.2 DNA precipitation from aqueous solution

Ethanol precipitation The DNA solution was mixed with 0.1 volume of 3M sodium acetate (pH 5) and 2.5 volumes of absolute ethanol and precipitated at -70°C for at least 30 minutes. This was then spun in a microcentrifuge at 15,000 r.p.m. for 15 minutes. The supernatant was discarded, and the pellet washed in 70% ethanol (-20°C) by vortexing. The suspension was re-centrifuged as above, the supernatant again discarded, and the pellet dried under vacuum. The above procedure was also employed to precipitate total RNA preparations. The dried DNA pellet was then resuspended as required in a suitable volume of water or TE buffer (with added Ribonuclease Cocktail if needed).

Isopropanol precipitation As for ethanol precipitation, except only 1 volume of isopropanol was used, thus keeping the total volume smaller.

2.5.3 Determination of DNA and RNA concentrations

DNA and RNA concentrations were determined by measuring the absorption of diluted solutions at 260 nm. An OD_{260} of 1.0 represents a DNA concentration of 50 $\mu g\ ml^{-1}$ for double-stranded DNA, or 40 $\mu g\ ml^{-1}$ for single-stranded RNA.

DNA purity was determined by measuring absorption at 260 and 280 nm. Pure double-stranded DNA should give a 260/280 ratio close to 1.8 and single-stranded RNA should give a ratio nearer to 2.0.

2.5.4 Small-scale plasmid preparation

Routine preparations of plasmid DNA were performed using a modification of the alkaline lysis method of Birnboim and Doly (1979). 5 ml of LB broth (containing antibiotics for selection) were inoculated with a single colony of the plasmid-bearing strain, and shaken overnight at the appropriate temperature (typically 37°C). The overnight culture was centrifuged at 4000 r.p.m. for 10 minutes in a bench-top centrifuge. The supernatant was discarded, and the bacterial pellet resuspended in 0.1 ml of buffer containing 1% glucose, 10mM EDTA, and 25mM Tris–HCl (pH 8.0). To this cell

suspension 0.2 ml of 0.2M NaOH containing 1% SDS was added, the contents were mixed by inversion of the tube and placed on ice for 5 minutes. 150 µl of 3 M Na acetate (pH 5.0) were then added, the solution vortexed, and left on ice for a further 5 minutes. The mixture was spun in a microcentrifuge for 10 minutes in order to pellet the precipitated chromosomal DNA and insoluble cellular debris. The resulting supernatant (~0.4 ml) was transferred to a fresh microcentrifuge tube and 1 volume of phenol:chloroform (phenol saturated with TE (pH 8.0) plus an equal volume of chloroform) was added, vortexed, and spun in a microcentrifuge for 2 minutes. The upper aqueous phase was transferred to a fresh tube and the plasmid DNA could then be recovered from solution by ethanol precipitation. In this case no extra salt was required in order for precipitation to occur. Typically the final pellet of nucleic acid was resuspended in TE buffer containing Ribonuclease Cocktail (from HT Biotechnology Ltd; RNase A (1 mg ml^{-1}), RNase T1 (20,000 U ml^{-1})). A 5 ml overnight culture typically yielded about 3–5 µg of plasmid DNA.

2.5.5 Restriction of DNA

Endonuclease cutting of DNA was typically performed in volumes of between 20 and 100 µl. These contained the requisite amount of DNA (usually 0.5–2 µg) and the appropriate restriction buffer at 1x concentration. The restriction enzyme was typically used at a two- to five-fold excess, i.e. 2–5 units per µg of DNA. The digests were made up to their final volume using distilled water. The complete restriction digests were incubated at the recommended temperature (usually 37°C) for 1–2 hours. The products of the reaction were either directly analysed by agarose gel electrophoresis, or phenol extracted, ethanol precipitated and dissolved in a suitable volume of TE buffer for further manipulations.

Partial digestion of DNA. For partial digestion of DNA, ten two-fold serial dilutions of restriction enzyme were added to fixed amounts of DNA, with 0.5 units of enzyme μg^{-1} DNA representing the highest enzyme:DNA ratio. The digests were incubated at the appropriate temperature for 1 hour and the reaction terminated by the addition of loading buffer (see 2.5.8). The products of the reactions could then be analysed by agarose gel electrophoresis.

2.5.6 Ligation of DNA

Ligations of DNA were typically performed in a final volume of 10 µl. These contained between 0.5–1 µg total DNA (with insert DNA in a 2- to 20-fold molar excess over the vector DNA), 1x ligation buffer and T4 DNA ligase. 0.2 Weiss units of ligase were used for the ligation of cohesive DNA termini, and 1 unit of the enzyme for the ligation of blunt-ended molecules. The reactions were incubated overnight at 16°C. Between 1 and3 µl of the reaction mixture were then used to transform competent cells of an appropriate strain of *E. coli.*

2.5.7 'Filling in' of recessed 3' termini

Klenow enzyme was used to fill in the recessed 3' termini generated by various restriction enzymes to give blunt-ended DNA molecules. Reactions were performed in a final volume of 30 µl containing 1 µg DNA, 1x Klenow buffer, all four dNTPs each at a concentration of 33 µM and 1 unit of Klenow enzyme. The tubes were incubated for 15 minutes at 25°C. The reactions were stopped and the unincorporated nucleotides removed by increasing the reaction volume to 200 µl with TE, phenol extracting and ethanol precipitating the DNA.

2.5.8 Agarose gel electrophoresis

Size fractionation of DNA fragments was carried out by agarose gel electrophoresis. The gels were made up by melting the appropriate amount of agarose (usually between 0.8 and 1.5%) in 1x TAE buffer. Samples, typically containing 100 to 500 ng of DNA in 1x loading buffer, were applied to the wells at one end of the gel. Electrophoretic separation was carried out, with the gels just immersed in 1x TAE buffer, at 20 to 70 mA constant current. After completion of electrophoresis, gels were stained in 1x TAE buffer containing 0.5 µg ml^{-1} ethidium bromide for about 30 minutes, then destained in distilled water for about 45 minutes. The gels were then photographed using Polaroid film and UV transillumination. The smallest amount of DNA detectable with ethidium bromide is about 2 ng. More than 500 ng of DNA per well will cause smearing.

2.5.9 Isolation of DNA from agarose gel slices

DNA was isolated from agarose gels by employing either a modification of the proprietary Geneclean II method, or by centrifugation of the DNA through siliconised glasswool.

Geneclean II-like procedure This utilises a silica matrix which binds DNA only in high-salt solutions. The appropriate DNA band was located in an ethidium bromide-stained gel under UV transillumination and excised using a scalpel blade in as small a volume of agarose as possible. Gel slices were transferred to microcentrifuge tubes, the weight of the slice determined and 3 volumes (0.1 g equals ~100 µl) of 6 M NaI added. These were incubated at 55°C for 5 minutes or until the gel slice had dissolved. 10 µl of silica (Sigma S-5631; 100 mg ml^{-1} in 3 M NaI) were added, the suspension mixed and put on ice for 5 minutes. Pellets were washed twice by briefly spinning in a microcentrifuge for 5 seconds followed by resuspension in 0.7 ml of wash buffer (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 2.5 mM EDTA, 50% v/v ethanol). After the final wash, the pellets were allowed to air-dry. The DNA was eluted from the silica matrix by resuspending the pellet in 11 µl of water for 5 minutes at 55°C. After centrifugation for 30 seconds, the DNA-containing supernatant was transferred to a fresh tube.

Siliconised glasswool method. The band containing the desired DNA fragment was cut out of a gel as described above. A 0.5 ml microcentrifuge tube was punctured at the bottom with a narrow-bore needle and half filled with siliconised glasswool. Care was taken not to pack this too tightly. The gel slice was placed on top of the glasswool and the 0.5 ml microcentrifuge tube was placed inside a 1.5 ml microcentrifuge tube. The tubes were spun at 7000 rpm in a microcentrifuge for 10 minutes, resulting in separation of the DNA solution (which was collected in the larger tube) from the agarose (which was retained on the glasswool). The DNA solution was transferred to a fresh microcentrifuge tube, extracted with phenol:chloroform and ethanol precipitated before further use.

2.5.10 Preparation of bacteriophage M13 DNA

5 ml of LB broth were inoculated with 50 µl from an overnight culture of TG1. To this either 100 µl of an M13 phage suspension, (about 1/10 of a single plaque), or an entire M13 plaque from an agar plate was added. This culture was shaken at 37°C for about 5 hours, and then centrifuged at 4,000 r.p.m. for 10 minutes. The resulting bacterial pellet could be used to prepare the double-stranded replicative form of M13 DNA, and the supernatant used to prepare single-stranded M13 DNA (or used as a fresh bacteriophage suspension).

Preparation of double-stranded M13 DNA. The bacterial pellet was washed once in bacterial buffer and the double-stranded DNA isolated using the small-scale method for plasmid DNA isolation (2.5.4).

Preparation of single-stranded M13 DNA. 1.3 ml of the bacteriophage suspension were transferred to a fresh microcentrifuge tube. The phage were precipitated on ice for at least 30 minutes by adding 200 µl of 20% polyethylene glycol (PEG8000) in 2.5 M sodium chloride. The precipitated phage were recovered by spinning twice in a microcentrifuge for 5 minutes and discarding the supernatant. The bacteriophage pellet was resuspended in 100 µl TE with vigorous vortexing. 1 volume of phenol (equilibrated with Tris–HCl (pH 8.0)) was added and the suspension mixed by vortexing for 30 sec.This mixture was allowed to stand at room temperature for 5 minutes, vortexed again, and spun for 5 minutes. The upper aqueous phase was removed and placed in a fresh tube. The volume of the sample was adjusted to 0.3 ml and the single-stranded DNA recovered by ethanol precipitation. Using this method the yield of single-stranded DNA was approximately 5–10 µg DNA per ml of infected culture and was of sufficient quality for both DNA sequencing and site-directed mutagenesis reactions.

2.5.11 Site-directed mutagenesis

Specific base changes in manipulated nucleotide sequences were effected by employing an adaptation of the procedure described by Kunkel (1985). This method of mutagenesis without phenotypic selection augments the yield of mutants, by taking advantage of a strain of *E. coli* that is deficient in two enzymes, deoxyuridine triphosphatase and uracil *N*-glycosylase. These

enzymes are encoded by the genes *dut* and *ung* respectively. The combined effect of these two deficiencies is that the cells frequently incorporate uracil bases instead of thymine bases into their DNA (because of the increased intracellular pool of dUTP) and are unable to repair the mistake. Single-stranded M13 template DNA prepared from these cells will contain, on average, four to six uracil residues per thousand bases. During *in vitro* complementary strand synthesis primed by the mutagenic oligonucleotide, thymine bases are incorporated correctly into the nascent strand, producing a heteroduplex product in which only one strand (the wild-type) contains uracil residues. When the heteroduplex molecule is reintroduced into an *ung*$^+$ strain that possess a functional uracil *N*-glycosylase, uracil residues are rapidly removed from the DNA, generating sites which block replication and are susceptible to cleavage by specific endonucleases. This activity causes the template strand to be preferentially degraded, leaving intact the *in vitro* synthesised strand containing the introduced mutation. A concomitant reduction in wild-type M13 infectivity (of approximately 10^5 fold) gives a high enough yield of mutants (greater than 50%) to permit direct DNA sequencing as the method for identifying mutants.

Template DNA is prepared by repeatedly growing a recombinant M13 phage which carries the target sequence, in strain CJ236 (*dut*, *ung*). Usually two to three passages through CJ236 gave M13 phages with a 10^5–10^6 fold reduced infectivity in an *ung*$^+$ strain. ssDNA was used as a template to which an oligonucleotide containing the desired mutation was annealed. Mutant oligonucleotides (approximately 20 bases in length incorporating 1-3 nucleotide substitutions) were phosphorylated by mixing 200 pmoles of oligonucleotide in 10 µl of water with 4.5 units of bacteriophage T4 polynucleotide kinase in 0.1 M Tris-HCl (pH 8.0), 0.01 M $MgCl_2$, 5 mM dithiothreitol, and 0.5 mM ATP. The reaction was incubated for 45 minutes at 37°C. The enzyme was then inactivated by heating for 10 minutes at 65°C.

The oligonucleotide was annealed to the uracil-enriched template DNA by mixing 0.5 pmole of ssDNA with 10 pmoles of oligonucleotide in 20 mM Tris-HCl (pH 7.7), 2 mM $MgCl_2$ and 50 mM NaCl. This was placed in a beaker of water at 70°C and allowed to cool to 30°C. The tube was then placed on ice.

The oligonucleotide was used as a primer for the *in vitro* synthesis of a uracil-free DNA strand complementary to the template strand. To the annealing mixture (20 µl) were added 1 Weiss unit of bacteriophage T4 DNA ligase, 1 unit of T4 DNA polymerase and each of the four dNTPs at a final concentration of 0.5 mM. These were incubated, at a final volume of 24 µl, in 10 mM Tris-HCl (pH 7.7), 5 mM $MgCl_2$, 2 mM dithioerythritol and 1 mM ATP for 90 minutes at 37°C. The tube was placed on ice and 80 µl of 10 mM Tris-HCl (pH 8.0), 10mM EDTA were added to stop the reaction.

The product of the reaction was used to transfect an M13-sensitive *ung*$^+$ strain of *E. coli* (generally TG1; see Section 2.3.1). The plaques obtained were screened for incorporation of the mutation by sequencing ssDNA.

2.5.12 DNA sequencing

Introduction. DNA sequencing was performed using the Pharmacia T7 Sequencing Kit. The protocol employed is an adaptation of the chain-terminating dideoxynucleotide sequencing method developed by Sanger *et al* (1977). T7 DNA polymerase replaces the Klenow fragment of *E. coli* DNA polymerase I as the enzyme that catalyses primer extension. This has the advantage of creating longer chain-terminated fragments with a more even distribution of label between fragments. The major practical difference in using T7 DNA polymerase is that the primer extension reactions are performed in two stages, a labelling reaction and a termination reaction. The two stages are required because the enzyme uses dideoxynucleotides very readily, and therefore in order to allow the synthesis of long chain-terminated fragments, dideoxynucleotides are excluded from the first stage of the reaction. Even so, the time required for the reactions using the T7 enzyme is considerably less than those using Klenow enzyme.

Annealing of primer to single-stranded template. The DNA templates used in the sequencing reactions were all single-stranded M13 DNAs and were purified as indicated previously (2.5.10). The concentration of the template was adjusted to 1 µg μl^{-1} in TE. The oligonucleotide was adjusted to a concentration of 0.80 µM.

The following were added to a microcentrifuge tube on ice:

template DNA (1μg μl^{-1})	2 μl
primer (0.80 μM)	2 μl
annealing buffer	2 μl
water	8 μl
Total	14 μl

The contents of the tube were mixed and incubated for 10 minutes at 60°C. The denatured template/oligo mix was then left to anneal for at least 10 minutes at room temperature.

Sequencing reaction. For each template to be sequenced, four microcentrifuge tubes or wells of a microtitre plate were labelled 'A', 'C', 'G' and 'T' and 2.5 μl of the corresponding dideoxynucleotide mix added to each tube or well. To the tube containing the annealed template and primer the labelling mix, (dCTP, dGTP and dTTP in solution), T7 DNA polymerase and labelled dATP were added as follows:

Annealed template and primer	14 μl	
Labelling mix	3 μl	
[α-^{35}S] dATPαS	1 μl	(=10 μCi)
Diluted T7 DNA polymerase (1.5 units μl^{-1})	2 μl	

This labelling reaction was incubated at room temperature for 5 minutes. While this was proceeding the previously dispensed sequencing mixes were incubated in a water bath for one minute at 37°C. After the 5 minute incubation of the labelling reaction, 4.5 μl was added to each of the prewarmed sequencing mixes and returned to the water bath for a further 5 minutes to allow chain-termination to occur. Finally, 5 μl of Stop Solution (98% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue) were added to each reaction, which could then be stored at –20°C until required for electrophoresis. Prior to loading onto the sequencing gel, the samples were heated for 2 minutes at 80°C to denature the DNA. Immediately after denaturation, 1.5 to 2.5 μl of each sample was loaded onto the gel.

Electrophoresis of DNA sequencing reactions The products of dideoxynucleotide chain-termination were electrophoresed through an 8% polyacrylamide denaturing gel on a 30 x 40 cm BRL sequencing apparatus. The glass

sequencing gel plates were cleaned with ethanol and chloroform, assembled using 0.2 mm spacers and taped together.
The gel was prepared by adding together the following:

40% (w/v) bis-acrylamide (filtered)	20 ml
Urea	50 g
Water	30 ml
10x TBE	10 ml

Once dissolved, 1 ml of a 10% ammonium persulphate solution was added followed by 35 µl of TEMED. After mixing, the solution was poured between the sequencing plates. The flat edge of a 60-well shark-tooth comb was pushed between the plates to create a space for the wells at the top of the gel. Clingfilm was wrapped round the exposed areas of the gel and each edge of the gel was clamped with bulldog clips. The gel was then set aside for at least 1 hour to allow polymerisation. Once set, the bulldog clips, tape and comb were removed and distilled water was washed along the top of the gel. The shark-tooth comb was then inverted with the points downwards just touching the surface of the gel. It was then clamped into the sequencing apparatus and 1x TBE solution poured into the top and bottom reservoirs. The gel was then pre-run at 1200-1500 V for 30-60 minutes. After this the gel was ready to be loaded with the sequencing reactions. The samples were loaded in the order A, C, G and T immediately after denaturing the DNA (see above). The gel was then electrophoresed at 1200-1500 V (approximately 30 mA, 43 W) until the bromophenol blue dye-front reached the bottom of the gel. Once electrophoresis was complete, the glass plates were removed from the apparatus and separated. The gel was fixed in 10% (v/v) methanol and 10% (v/v) acetic acid for 20 minutes in order to remove excess urea. A damp sheet of blotting paper was placed on top of the gel followed by a sheet of dry blotting paper. Even pressure was applied and the papers were peeled off the glass plate taking the gel with them. The gel and paper sandwich was then dried in a vacuum gel drier for 1 hour at 80°C. Cronex X-ray film was exposed overnight to the dried gel in an autoradiography cassette at room temperature.

2.5.13 Amplification of DNA using the Polymerase-Chain-Reaction

Specific regions of plasmid DNA were amplified by the Polymerase-Chain-Reaction (PCR) in order to confirm the presence and orientation of certain fragments believed to be ligated to the target vector. The source of the plasmid template was either purified mini-prep DNA or whole cells taken directly from the transformation plate (Colony-PCR). A master mix for multiple reactions was usually made, with the following final composition per reaction:

1x Reaction buffer
1.5 mM $MgCl_2$
20-200 µM each dNTP
100 µg ml^{-1} BSA
0.2 µM oligonucleotide primer 1
0.2 µM oligonucleotide primer 2
2 units/reaction *Taq* DNA polymerase
100 ng DNA template
Distilled water to 100 µl

The cycle programme was determined by the expected length of the product and the GC content of the primer. A typical sequence was:

Denaturation:	1 minute at 95°C
Annealing:	1 minute at the primer T_m -5°C
Extension:	1 minute at 72°C (1 minute kb^{-1} of PCR target)
No. of cycles:	30

The melting temperature (T_m) of the primers was calculated with the following formula:

$$T_m\,(°C) = 2(N_A + N_T) + 4(N_G + N_C)$$

where N equals the number of each base in the primer.

Colony-PCR differed only in respect of template preparation and denaturation. A single transformant was picked off a plate with a sterile toothpick and streaked on the appropriate selection medium. This same toothpick was then dipped into the PCR tube containing the reaction mix

and twisted several times. An extra one-off denaturation step of 3 minutes at 95°C was carried out at the start of the amplification sequence to ensure complete plasmid liberation from the bacterial cells.

2.6 RNA techniques

2.6.1 Total prokaryotic RNA preparation

Note: Appropriate steps were taken to avoid ribonuclease contamination during the manipulation of RNA i.e. always wearing gloves, soaking glassware in freshly prepared 0.1% (v/v) diethyl pyrocarbonate (DEPC) and using sterile, disposable plasticware. All solutions were treated with 0.1% DEPC or made up with 0.1% DEPC-treated water where necessary (left at room temperature overnight and autoclaved to remove any trace of DEPC).

RNA preparations were performed as described by He *et al* (1993) and Blomberg *et al* (1990). A fresh overnight culture (inoculated from a single colony) of a bacterial strain was diluted 100-fold into fresh LB broth (plus antibiotics where appropriate) and shaken at 37°C until the OD_{600} reached about 0.5. When plasmid-containing strains were to be induced, IPTG was added at 0.2 mM and incubation continued for a further hour. A 5 ml aliquot of the culture was added to 1 ml of stop solution (5% phenol in 95 % ethanol) at room temperature to prevent enzymatic degradation of the metabolically unstable RNA during further preparation. After the sample was centrifuged for 5 minutes, the pellet was resuspended in 0.25 ml of 0.36% glucose and transferred to a microcentrifuge tube. The cells were lysed by boiling for 15 seconds after the addition of 0.5 ml of lysing solution (1% SDS, 0.2 M NaCl, 20 mM Na_2EDTA) at 100°C and vigorous mixing for 10 seconds. The RNA was isolated by extracting the samples (a cleared lysate) twice with an equal volume of 60°C phenol (equilibrated with 10 mM Tris-HCl pH 7.5, 1 mM Na_2EDTA), followed by one extraction with chloroform. After ethanol precipitation and washing, the samples were resuspended in 45 µl of water. Contaminating DNA was removed using RQ1 RNase-free DNase I (Promega). 5 µl of RQ1 10x buffer (400 mM Tris-Hcl, pH 7.9, 100 mM NaCl, 60 mM $MgCl_2$, 100 mM $CaCl_2$) and 1 µl of RQ1 DNase (1 U μl^{-1}) were added to the sample and incubated at 37°C for 15 minutes. After inactivation of the enzyme by phenol extraction and ethanol precipitation, the purified RNA was resuspended in 5 µl of water. The concentration of RNA was determined by optical density (an OD_{260} of 1.0 represents 40 µg ml^{-1}). To prevent degradation of the RNA, 0.5 µl of RNasin® Ribonuclease inhibitor (40 U μl^{-1}) was added to the preparation. The samples were stored at -20°C.

2.6.2 End-labelling (5') of oligonucleotide primers

End-labelling of oligonucleotides was necessary for visualisation of primer extension products. 10 pmol of oligonucleotide in 2 µl of water was mixed with 1 µl of T4 Polynucleotide Kinase 10x buffer (500 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 50 mM DTT, 1 mM spermidine), 3 µl of [γ-^{32}P] ATP (3,000 Ci mmol^{-1}, 10 mCi ml^{-1}), 1µl of T4 Polynucleotide Kinase (8-10 U µl^{-1}) and water to a final volume of 10 µl. The reaction was incubated at 37°C for 30 minutes followed by inactivation of the enzyme at 90°C for 2 minutes. The final concentration of the primer was brought to 100 fmol µl^{-1} by adding 90 µl of water. The labelled oligonucleotide was separated from unincorporated [γ-^{32}P] ATP by passing it through a NAP-5 column (Pharmacia). The column was equilibrated with 10 ml of water without allowing it to dry out. The labelled oligonucleotide was run into the column followed by 200 µl aliquots of water. As each aliquot passed through the column, it was collected in a microcentrifuge tube. Fractions containing the labelled oligonucleotide appeared in the first radioactive peak, identified by monitoring with a Geiger counter. Typically this was detected within the first six to eight fractions. The second peak, which usually overlapped the first one, contained the unincorporated label. The fractions containing the labelled oligonucleotide were pooled and stored at -20°C.

2.6.3 Primer extension analysis

Primer extension analysis was used to locate the 5'-end of a specific mRNA. An end-labelled oligonucleotide was hybridized to the target RNA (preferably about 100 nucleotides from its 5' end) and utilised as a primer by reverse transcriptase in the presence of deoxynucleotides. The resultant cDNA was analysed on a denaturing polyacrylamide gel. Annealing of the primer to the specific mRNA was carried out by adding 5 µl of AMV primer extension 2x buffer (100 mM Tris-HCl pH 8.3 at 42°C, 100 mM KCl, 20 mM $MgCl_2$, 20 mM DTT, 2mM each dNTP, 1 mM spermidine) and 1 µl (100 fmol) of the ^{32}P-labelled primer to the 5 µl sample of total RNA. The reaction was incubated at the melting temperature (T_m) for the primer for 20 minutes and then placed at room temperature to cool for 10 minutes. Extension of the annealed primer was performed by adding 9 µl of extension mix (5 µl of AMV primer extension 2x buffer, 1.4 µl of 40 mM sodium pyrophosphate, 1 µl of AMV reverse transcriptase (5-10 U µl^{-1}) and water to a final volume of 9

µl) to the tube containing the annealed primer/RNA and incubating at 42°C for 30 minutes. The reaction was stopped by adding 20 µl of loading dye (98% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue). 20 µl of the 40 µl volume (the remaining sample was stored at -20°C) was heated to 90°C for 10 minutes and immediately loaded on a denaturing polyacrylamide gel (8% acrylamide, 7M urea, and TBE 1x buffer) that was prepared and processed as indicated for a sequencing gel. An [α-^{35}S] dATPαS-labelled sequencing reaction from a known template/primer pair was used as the fragment size marker.

2.7 Protein techniques

2.7.1 β-galactosidase enzyme assay

The level of expression of an exogenously introduced copy of the *lacZ* gene under foreign promoter control was determined by assaying the resultant β-galactosidase activity. The method used is essentially that of Miller (1972). When assays were carried out in strains possessing a chromosomal copy of the *lacZ* gene, glucose at 0.2% was added to the medium to minimise transcription from this source (catabolite repression). Chromosomally encoded *lacZ* expression was undetectable under these conditions. Otherwise, cells were grown in the media indicated in the text. A fresh overnight culture (inoculated from a single colony) was diluted into fresh LB broth (plus antibiotics where appropriate) at a low optical density and grown for at least 3 hours prior to sampling. Early exponential cultures were maintained in steady state at an OD_{600} between 0.05 and 0.3 by re-inoculating into prewarmed broth at 37°C as necessary.

All assays were carried out in small glass test-tubes. Samples of either 0.5 ml or 0.1 ml (depending on the expected level of β-galactosidase activity) were taken and added to 0.5 ml or 0.9 ml respectively of Z-buffer (0.06 M Na_2HPO_4, 0.04 M NaH_2PO_4, 0.01 M KCl, 0.001 M $MgSO_4$, 0.05 M β-mercaptoethanol and 0.005% SDS). The cells were permeabilised by the SDS present in the Z-buffer and the addition of 15 µl of chloroform to each sample. The tubes were vortexed for 10 seconds immediately after sampling. At this stage they could be sealed and stored overnight at 4°C.

The assay was performed by adding 0.2 ml of 4 mg ml^{-1} ONPG (o-nitrophenyl-β-D-galactoside) in 0.1 M MOPS (pH 7.0) to each sample. These were incubated at 30°C until sufficient yellow colour (o-nitrophenol) had appeared such that when 0.5 ml of Na_2CO_3 was added to stop the reaction, the absorbance at 420 nm was between 0.1 and 2.0. Standard curves constructed in this laboratory by T. Paterson (unpublished observation) had demonstrated that the absorbance measurements are linear within this range. Samples were measured at both 420 nm (o-nitrophenol) and at 550 nm (to correct for light scattering by cell debris). In cases where samples had been taken at high cell-density, the samples were spun in a microcentrifuge after

incubation to remove cell debris; in this case only the absorbance at 420 nm was determined.

Total enzyme activity was calculated as:

$$1000x\ (OD_{420}-1.75xOD_{550})/(TxV)$$

where T represents the incubation time in minutes and V represents the volume of the sample in millilitres (generally 0.5 or 0.1).
The specific activity is defined as total enzyme activity divided by the OD_{600} of the culture at the time of sampling (i.e. enzyme units per total cell-mass) and is expressed in Miller Units (MU).

2.7.2 *In vitro* DNA-directed protein translation using circularised plasmid DNA

Cell-free coupled transcription/translation of cloned DNA sequences was carried out *in vitro* using the *E. coli* S30 extract system for circular DNA from Promega (Product # L1020). The procedure, a modification of the method described by Zubay (1973), employs an S30 extract prepared from an *E. coli* strain B that is deficient in specific protease activity and is free of any amber suppressor mutations. The S30 extract is a fraction that includes RNA polymerase, 70S ribosomes and other protein factors necessary for translation. A premix lacking amino acids is the source for other essential requirements including ribonucleotide triphosphates, tRNAs, phosphoenol pyruvate and appropriate salts. Amino acids are supplied in an equimolar mixture lacking in either methionine or leucine. These are added to the reaction in a radio-labelled form, allowing visualisation of the translated products by autoradiography.

A master mix for multiple reactions was usually made, with the following composition per reaction:

Amino acid mixture minus methionine	5 µl
S30 Premix without amino acids	20 µl
[^{35}S] methionine (1,200 Ci mmole^{-1} at 15 mCi ml^{-1})	1 µl
S30 Extract, circular	15 µl
RNasin® Ribonuclease Inhibitor (40 U µl^{-1})	1 µl
Nuclease free water	to final volume of 50 µl

DNA template was added to an aliquot of the reaction mixture at 4 µg per 50µl. The reaction was incubated at 37°C for 1-2 hours and stopped by chilling on ice for 5 minutes. The samples were then prepared for analysis by SDS-PAGE. A 5 µl aliquot from each reaction was added to 20 µl of acetone (to remove PEG from the extract) and chilled on ice for 15 minutes. The samples were spun in a microcentrifuge for 5 minutes and the supernatant discarded. After drying under vacuum for 15 minutes, the samples were resuspended in 20 µl of 1x SDS-PAGE loading buffer. Immediately prior to loading on a 12% SDS-PAGE gel, an aliquot (sufficient for 10 µl per lane) of the sample was heated at 100°C for 2-5 minutes. The remaining sample was stored at -20°C. After running, staining and drying the gel (section 2.7.5), Cronex X-ray film was exposed overnight at room temperature to the dried gel. In some instances, (when using [^{3}H] leucine as the radio-label) the translation products were detected by fluorography, which is more sensitive than autoradiography. In this case, the destained gel was infused with Amplify™ (Amersham), an organic scintillant, for 30 minutes prior to drying the gel. Pre-flashed Cronex X-ray film was exposed to the dried gel for an appropriate time at -70°C, in an X-ray cassette with an intensifying screen.

2.7.3 Overexpression of proteins *in vivo* by T7 RNA polymerase

The overexpression of a protein was carried out using the T7 system (Studier and Moffat, 1986; Tabor and Richardson, 1985). The gene encoding the target protein was cloned immediately downstream from the T7 RNA polymerase promoter (ϕ10) in a pT7 vector. This was transformed into a strain (BL21(DE3)/pLysS) that has been lysogenised with a λ phage encoding T7 RNA polymerase under the inducible control of the *lacUV5* promoter. Expression of T7 RNA polymerase (and hence the cloned protein) was induced by treatment with IPTG. Strain BL21 (an *E. coli* B strain) was used because it is deficient in both *lon* and *ompT* proteases. The *lacUV5* promoter is somewhat leaky. Plasmid pLysS (a derivative of pACYC184) reduces the system's ability to transcribe target genes in uninduced cells because it encodes T7 lysozyme, a natural inhibitor of T7 RNA polymerase. A fresh overnight culture of BL21(DE3)/pLysS harbouring a pT7 vector encoding the gene to be overexpressed (and a second culture bearing the parental pT7 plasmid) were spun down and resuspended in 1 ml of fresh LB broth. 0.5 ml of these cultures were inoculated into 24.5 ml of Spizizen minimal medium (300 ml water, 80 ml 5x Spizizen Minimal Salts (10 g $(NH_4)_2$ SO_4, 70 g

K_2HPO_4, 30g KH_2PO_4, 5g Tri-Sodium Citrate, 1 g $MgSO_4$, distilled water to 1 litre), 10 ml glucose (20% w/v), 0.5 ml thiamine/vitamin B_1 (1 mg ml^{-1}), 0.2 ml ampicillin (100 mg ml^{-1}), 0.4 ml chloramphenicol (20 mg ml^{-1})) and grown at 37°C until an OD_{600} of 0.6-0.8 was reached. Four aliquots of 0.5 ml were removed from each culture and 3 µl of IPTG (20 mg ml^{-1}; final conc. 0.5 mM) were added to three of them (the fourth aliquot was an uninduced control). The samples were incubated at 37°C. At one, two and three hour intervals post-induction, an aliquot was chilled on ice. The cells were pelleted in a microcentrifuge and resuspended in 100 µl of 1x sample buffer by vortexing. The tubes were boiled for 5 minutes immediately prior to loading a 30 µl sample onto a polyacrylamide gel.

2.7.4 *In vivo* radiolabelling of proteins that are overexpressed by T7 RNA polymerase

The sensitivity and specificity of the system described above (section 2.7.3) was increased by radiolabelling nascent proteins in the presence of rifampicin, an antibiotic that blocks transcription initiation from native promoters by inhibiting the activity of *E. coli* RNA polymerase. Cultures were prepared and grown in the manner described in section 2.7.3. Four aliquots of 0.5 ml were removed from each culture and 3 µl of IPTG (20 mg ml^{-1}; final conc. 0.5 mM) were added to two of them. After incubating all four aliquots at 37°C for 30 minutes, 1 µl of rifampicin (100 mg ml^{-1}; final conc. 200 µg ml^{-1}) was added to one IPTG-treated sample and one non-IPTG-treated sample. All four aliquots were incubated at 37°C for a further 45 minutes. Each sample was pulse-labelled for 5 minutes with 5 µCi of labelled amino acid. After chilling on ice, the samples were treated in the manner described in section 2.7.3.

2.7.5 Polyacrylamide gel electrophoresis of proteins

Protein resolution and visualisation were carried out using SDS–polyacrylamide gel electrophoresis with a discontinuous buffer system (Laemmli, 1970). In virtually all cases a 12% resolving gel and 5% stacking gel were employed. For the rapid analysis of proteins a mini–gel apparatus (Hoeffer Scientific Instruments 'Mighty Small' SE 250) was used. This used 10 x 5 cm gels, which could be electrophoresed in about 45 minutes. For improved resolution and better quality gels, a larger apparatus (Hoeffer SE

600) was employed with a gel size of 11 x 14 cm. See table 2.7.5 for the composition of typical gel solutions.

Table 2.7.5 Composition of a 12% polyacrylamide denaturing gel.

Component	12% Resolving gel (ml)	5% Stacking gel (ml)
Water	8.6	6
Acrylamide stock solution (40%)	6	1.25
4x Resolving gel buffer	5	-
4x Stacking gel buffer	-	2.5
10% SDS	0.2	0.1
10% Ammonium persulphate (fresh)	0.2	0.1
TEMED	0.008	0.01
Total	20	10

These solutions were made up on ice immediately prior to use, with the ammonium persulphate solution and the TEMED being added last. The resolving-gel solution was pipetted between the glass plates separated by 0.75 mm spacers; sufficient room was left for the stacking gel. Once the resolving gel had been poured it was layered with isobutanol saturated with 1x resolving-gel buffer and allowed to polymerize for 1 hour. The isobutanol was then discarded and the top of the gel was washed with distilled water. The stacking-gel solution was poured on top of the resolving gel, the comb inserted and polymerization allowed to occur. The comb was then removed and the wells washed out with 1x reservoir buffer, which was also used to fill up the buffer chambers of the apparatus. The samples could be loaded onto the gel at this stage.

Samples were mixed 1:1 with 2x SDS-PAGE-loading buffer, raised to 100°C for 3 minutes, and spun in a microcentrifuge for 5 minutes prior to loading. About 2–5 µl of sample per lane could be loaded onto the mini–gel and 10–30 µl on the larger apparatus. Typically, electrophoresis was carried out at a constant current of 15 mA in the stacking gel and 30 mA in the resolving gel,

until the bromophenol blue dye-front had run off the bottom of the gel. After electrophoresis, protein bands were visualised by staining with Coomassie blue dye. Gels were placed in staining solution for 15-30 minutes with gentle agitation. Dye that was not bound to protein was removed by transferring the gel to destaining solution for 2 to 24 hours. For preservation, the stained/destained gel was soaked in destaining solution plus 5% glycerol for 30 minutes and dried down on blotting paper using a vacuum gel drier at 60°C for 1 hour.

Table 2.7.6 Solutions used in SDS–PAGE

Acrylamide/Bis acrylamide stock solution	40% (w/v) acrylamide. Ratio 19:1 - Bis acrylamide. Filtered and stored at 4°C.	
4x stacking-gel buffer (0.5 M Tris-HCl)	15.25 g of Tris base dissolved in 200 ml distilled water, adjusted to pH 6.8 with concentrated HCl, made up to 250 ml, filtered and autoclaved.	
4x resolving-gel buffer (1.5 M Tris-HCl)	45.5 g of Tris base dissolved in 200 ml distilled water, adjusted to pH 8.8 with concentrated HCl, made up to 250 ml, filtered and autoclaved.	
10x reservoir buffer	30.2 g of Tris base and 144 g of glycine dissolved in 600 ml distilled water, made up to a final volume of1 litre and filtered. SDS was added to 0.1% in the final 1x buffer.	
2x PAGE loading buffer	4x stacking gel buffer	2.5 ml
	10% SDS	4 ml
	glycerol	2 ml
	β–mercaptoethanol	1 ml
	bromophenol blue	40 mg
	xylene cyanol	40 mg
	Distilled water	0.5 ml
	Total	10 ml
Staining solution	7% (v/v) acetic acid, 45% (v/v) methanol and 0.1% (w/v) Coomassie brilliant blue.	
Destaining solution	7.5% (v/v) acetic acid and 5% (v/v) methanol.	

CHAPTER 3

Characterisation of the *E. coli* poly(A) polymerase I promoter

3.1 Introduction

PcnB has been identified as an RNA processing enzyme with a poly(A) polymerase activity. RNase E and RNase III, the main endonucleases involved in the processing and decay of *E. coli* mRNA, have been shown to autoregulate their expression post-transcriptionally by destabilisation of their transcripts (Jain and Belasco 1995; Bardwell *et al*, 1989 respectively). Of the major exonucleases, PNPase determines the level of its expression by negative autoregulation at the level of translation, although only after its message has been cleaved by RNase III (Robert-Le Meur and Portier, 1992). It has also been demonstrated that RNase E and RNase III exert some control over the intracellular levels of PNPase (Takata *et al*, 1987; Hajnsdorf *et al*, 1994b) and RNase II (Zilhão *et al*, 1995). The modulation of RNase II expression by PNPase (Zilhão *et al*, 1996) further illustrates this network of incestuous regulation by the RNA processing enzymes. To determine whether PcnB controls the level of its own synthesis post-transcriptionally, an initial analysis was carried out to locate the 5' end of its transcript.

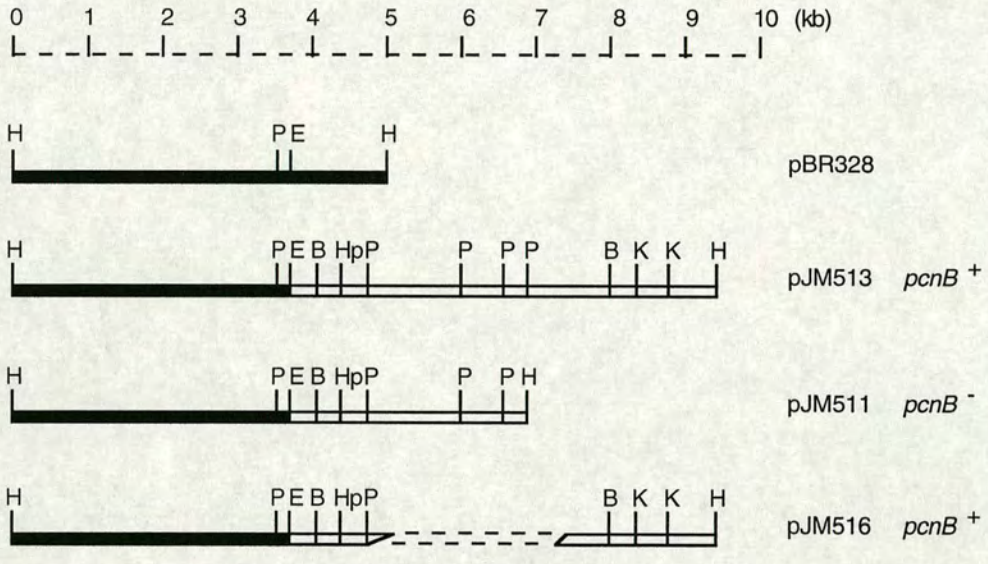
3.2 Identification of the *pcnB* promoter

Sequence analysis reported in the current literature to identify the promoter appeared inconclusive. Liu and Parkinson (1989) suggested that the -10 region begins approximately 140 nucleotides upstream from the inferred coding region. However, the N-terminal sequencing of PcnB by Cao and Sarkar (1992) revealed a translated product that extends 40 amino acids upstream of the previously inferred initiator methionine. This open reading frame continues upstream for a further 17 codons, spanning the previously proposed site of the promoter. Additional upstream sequence data obtained by these workers were considered by them to be devoid of any obvious transcription control elements.

3.2.1 The phenotypic difference between pJM516 and pJM513 defines the region of the *pcnB* promoter

By comparing the DNA content of two phenotypically different constructs previously made in this laboratory, each containing the complete *pcnB* coding sequence, I was able to predict the location of the *pcnB* promoter. The plasmid pJM513 encodes *pcnB* on a 5.3 kb fragment cloned from λ 115 of the Kohara library (March *et al*, 1989; Kohara *et al*, 1987). Diminished colony size is a characteristic of cells bearing this construct, consistent with the observation by Cao and Sarkar that excess PcnB and hence multiple copies of *pcnB* are detrimental to host viability (Cao and Sarkar, 1992). As part of the strategy to isolate a minimal fragment capable of complementing *pcnB21*, pJM513 had been subjected to deletion analysis. A partial digestion of pJM513 with *Pvu* II was followed by the addition of *Hin*d III or *Eco*R I linkers. Restriction of the individual plasmid products with *Eco*R I or *Hin*d III, as appropriate, yielded a series of deletions extending from one end or the other of the insert DNA to an internal *Pvu* II site. The constructs pJM511 (a *Hin*d III deletion) and pJM516 (an *Eco*R I deletion) were expected to be paired right- and left-hand side deletion products (respectively), originating from the same *Pvu* II site. Whereas the predicted host DNA content of pJM511 was confirmed, pJM516 had unexpectedly retained DNA originating from the left-hand end of the pJM513 insert (Figure 3.2.1).

Figure 3.2.1 Map of pJM513 and its deletion derivatives.



Creation of pJM516. DNA originating from pBR328 is shaded; DNA of chromosomal origin is unshaded; DNA deleted from pJM513 during construction of pJM516 is hyphenated; *pcnB*$^{+}$/$^{-}$ indicates ability or inability to complement *pcnB-21*; Restriction sites are: B, *BstE* II; E, *Eco*R I; H, *Hin*d III; Hp, *Hpa* I; K, *Kpn* I; P, *Pvu* II (Modified from March *et al*, 1989).

The ability to complement *pcnB21* is retained by pJM516, but without conferring the reduced colony size phenotype of pJM513. This behaviour suggests that the level of *pcnB* expression from pJM516 is lower than from pJM513 and could be explained by the loss of a positive regulatory element during the creation of pJM516.

3.2.2 Sequencing the pJM516 insert at the junction site of the left- and right-hand ends from pJM513

Because the pJM516 insert was the product of an aberrant cloning event (March *et al*, 1989), its precise nucleotide content was unknown. In order to determine whether a native *cis*-acting control element had been deleted from the pJM516 insert, I subcloned and sequenced a fragment of the insert that bridged the known ends and compared it with the sequence from the equivalent region in pJM513. The ~1.3 kb *Bst*E II fragment from pJM516 was end-filled with the Klenow fragment of DNA Polymerase I and cloned into the unique *Hinc* II site in M13BM21, a derivative of M13mp19 with an expanded polylinker, creating M13NB4 (Figure 3.2.2).

Figure 3.2.2 M13NB4.

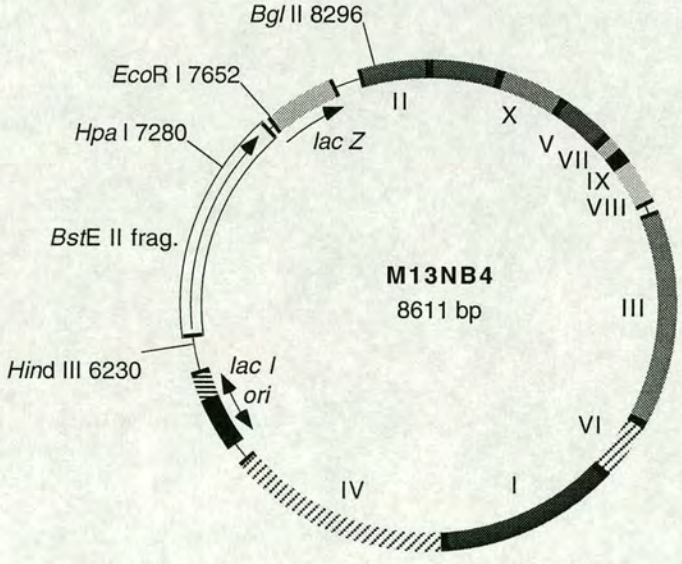


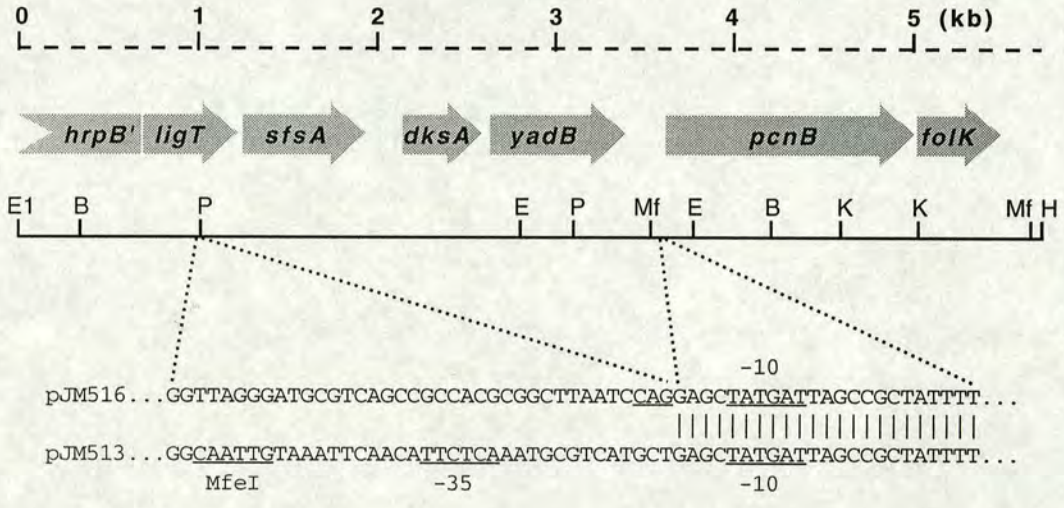
Figure 3.2.3. Sequence of M13NB4 insert.

Dideoxy-sequence ladder of the 1.3 kb *Bst*E II fragment from pJM516. Single-stranded template of M13NB4 DNA (bearing the *Bst*E II fragment) was sequenced from oligonucleotide N4212 (5′-d[GTCTGTTCTTTGCTATCG]-3′). The underlined triplet (CAG) indicates the 5′-half of the most upstream *Pvu* II site present in the pJM513 insert.

A C G T
A G T A T C G A G G A C C T A A T T
pJM516
junction

The two ends originating from the pJM513 insert were shown to juxtapose in pJM516 at the site shown in Figure 3.2.3. The sequence data from this gel are placed in context with the insert from pJM513 in Figure 3.2.4.

Figure 3.2.4 Sequence of pJM516 insert at the junction site of the left- and right-hand ends from pJM513.



Sequence data derived from M13NB4 and oligonucleotide N4212 (Figure 3.2.3). Shaded arrows indicate the coding region of genes present in the pJM513 insert and their direction of transcription. Expanding lines indicate the sequence found in pJM516 at the junction site of the left- and right-hand ends of the pJM513 insert. Below is shown the equivalent region from pJM513 extending from the right-hand end; the putative *pcnB* σ^{70} -35 region and the upstream *Mfe* I (Mf) site found to be deleted from pJM516 are underlined, as is the -10 region common to both inserts. The underlined triplet (CAG) indicates the 5'-half of the most upstream *Pvu* II site present in the pJM513 insert. Other restriction sites: B, *Bst*E II; E1, *Eco*R I; E, *Eco*R V; H, *Hin*d III; K, *Kpn* I; P, *Pvu* II. Map is drawn to scale.

3.2.3 The pJM516 insert lacks a recognisable -35 region

An examination of the sequence near the junction site revealed a plausible σ^{70} promoter sequence in pJM513; only the -10 element was retained in pJM516. For pJM516 to have retained the ability to complement *pcnB21*, transcription of *pcnB* must be initiated from a substitute site. There are three possible sources for such a site. Firstly, a weak hybrid promoter could have been created by the fusion of the left- and right-hand ends of the original pJM513 insert. However, an examination of the sequence around the site of

juxtaposition failed to reveal a surrogate -35 region. Alternatively, transcription could be directed from just the native -10 region; a reduced rate of transcription from such a non-canonical promoter could account for pJM516 not exhibiting the small colony phenotype associated with pJM513. Secondly, read-through transcription from the chloramphenicol acetyl transferase promoter in the back-bone vector, pBR328, could have been responsible for the expression of *pcnB* in pJM516 (the β-lactamase gene is transcribed in the reverse direction with respect to *pcnB* and the transcriptional initiation site for the tetracycline resistance gene was deleted during the construction of pJM513). I tested the possibility that the chloramphenicol promoter was responsible by blocking any possible read-through transcription from the back-bone vector with the bi-directional transcriptional/translational terminators present in the Ω fragment (Prentki and Krisch, 1984). This 2.0 kb fragment, from pHP45Ω, was cloned into the *Eco*R I site in pJM516 (at the junction between pBR328 DNA and the pJM516 insert) upstream of *pcnB*. The resultant construct, pJM516Ω, was still able to complement a *pcnB* strain as demonstrated by the level of single-cell ampicillin resistance. The third and most likely source for the transcriptional initiation of *pcnB* in pJM516 is from an upstream heterologous promoter, made proximal to the *pcnB* coding region by the loss of the intervening DNA. Two genes, *hrpB* and *ligT*, are located upstream from *pcnB* in pJM516. Both are incomplete and only *ligT* is transcribed in the same direction as *pcnB*. Since the 3'-end of *ligT* has been deleted along with all the intervening genes upstream from *pcnB*, any downstream transcriptional terminators will also have been removed. This would facilitate transcription of *pcnB* from a *ligT* promoter. Although at least one potential transcriptional initiation site was found in the sequence 5' to *ligT*, no further attempt was made to distinguish between this or the non-canonical promoter as the source of *pcnB* transcriptional initiation in pJM516.

3.2.4 Reduced colony size phenotype is associated with the presence of the -35 region

The sequence data from the pJM516 insert indicated that it might lack the *pcnB* -35 region. By comparing this sequence with the content of the pJM513 insert, it was found that the 2.1 kb *Mfe* I fragment (Figure 3.2.4) from pJM513 includes the putative -35 region with just 35 more nucleotides upstream from *pcnB* than is present in the pJM516 insert (ignoring the left-hand end from

pJM513). This additional DNA extends only 16 nucleotides upstream from the -35 region in the pJM513 insert. To demonstrate that the putative -35 region is part of the *pcnB* promoter (and also account for the absence of the reduced colony size phenotype with pJM516), I subcloned the 2.1 kb *Mfe* I fragment containing *pcnB*, from pJM513, into the compatible *Eco*R I site in pBR328. This insert extended from the *Mfe* I site indicated in Figure 3.2.4 to one 619 nucleotides 3′ of the *pcnB* stop codon. The small colony phenotype of host cells bearing the resultant 7 kb plasmid, pNB24, appears similar to that observed with pJM513. This result suggests that a critical *cis*-acting element not present in pJM516 resides within the additional upstream 35 nucleotides contained in the pNB24 insert.

3.2.5 Mapping the 5'-end of the *pcnB* message

The evidence derived from pJM516 seemed to indicate the approximate location of the putative *pcnB* promoter. With this information, I was able to map the 5′ terminus of the *pcnB* transcript by primer extension analysis. This technique depends upon the synthesis of a specific cDNA using an end-labelled primer complementary to the target RNA. The size of the cDNA reflects the distance from the primer to the 5′-end of the RNA. Total RNA was prepared from strain P90C harbouring a *pcnB′-lacZ* transcriptional fusion (either pNB14 or pNB15; see section 3.3 for a full description and construction details of these plasmids) or the vector backbone it was derived from (pASHOK). A labelled *pcnB*-specific oligonucleotide (D113; 5′-d[CGACTCGGGTAAAAATAGTACACCTCGGTA]-3′) was annealed towards the 5′-end of the mRNAs encoded by the *pcnB′-lacZ* fusion and the chromosomally derived copy of *pcnB*. The primer was extended by AMV reverse transcriptase and the extension endpoint was visualised by autoradiography after separation on a polyacrylamide gel (Figure 3.2.5). To determine the size of the primer extension endpoint and hence the location of the 5′-end of the *pcnB* transcript, a dideoxy-sequence ladder derived from single-stranded M13mp18 DNA sequenced with the Universal Primer (5′-d[GTAAAACGACGGCCAGT]-3′) was run alongside the extension products. The fragment size of the sequencing bands was determined by identifying the sequence that migrated the same distance as the primer extension

Figure 3.2.5 Identification of the 5' end of *pcnB* mRNA.

Mapping of the 5'-end of *pcnB* mRNA (from a *pcnB'-lacZ* fusion and native *pcnB* transcripts) by primer extension. Primer extension analysis was performed on total RNA prepared from P90C harbouring various plasmids. Reverse transcription was directed from oligonucleotide D113 (see text). pNB14 and pNB15 are products of the same cloning event that are believed to be identical. They encode a *pcnB'-lacZ* transcriptional fusion; pASHOK is the parent vector that lacks the fusion. The band marked P indicates the position of the primer extension endpoint. The dideoxy-sequence ladder was derived from single-stranded M13mp18 DNA sequenced with the Universal Primer and is therefore not related to *pcnB;* it serves only as a convenient nucleotide ladder. The numbers beside the sequencing ladder indicate the nucleotide co-ordinates of M13mp18 DNA from which the size of the bands were determined. Note that the sequence shown is of the template strand.

A C G T
pNB14
pNB15
pASHOK
1016
1004
998
P

endpoint and its location with respect to the Universal Primer binding site. The Universal Primer is complementary to the coding strand of M13mp18 DNA. It binds from nucleotide 6306 to 6290 (5'-3'). From the gel, the primer extension endpoint for *pcnB* mRNA is equivalent in size to nucleotide 6246 when M13mp18 is sequenced from the Universal Primer. Therefore, the size of the primer extension endpoint equals 6290 (the 3'-end of the Universal Primer) minus 6246 plus 17 (the Universal Primer is a 17mer) or 61 nucleotides. The primer extension oligonucleotide (D113) is complementary to the coding strand of *pcnB* mRNA. Based on co-ordinates from the pJM513 insert, it binds from nucleotide 3631 to 3602 (5'-3'). As the primer extension product is 61 nucleotides long, the first base to be transcribed from *pcnB* must be 3602 (the 3'-end of D113) minus 61 plus 30 (D113 is a 30 mer) or nucleotide 3571.

3.2.6 The *pcnB* -10 and -35 regions resemble a consensus σ^{70} promoter

A composite map of the transcriptional initiation region of *pcnB* was derived from the collective data (Figure 3.2.6).

Figure 3.2.6 Transcriptional initiation region of *pcnB*.

```
 3520       3530       3540       3550       3560       3570
...|.........|.........|.........|.........|.........|.........|.........
                                                      +1
                                                      ⇓
                      -35                      -10
ggcaattgtaaattcaacaTTCTCAaatgcgtcatgctgagcTATGATtagccgCtatttttt
                    ||  ||                   |||  ||
                    TTGACA   17 nt spacer    TATAAT      Consensus σ70
```

A composite map of the transcriptional initiation region of *pcnB*. Arrow indicates first transcribed base; hexanucleotides in uppercase are the inferred -10 and -35 sequences of a σ^{70} promoter; underlined bases are those with the highest percentage occurrence at that site; nt, nucleotide.

The *pcnB* promoter has a total raw score of 262 and a similarity score of 58.6 when analysed with a base-scores matrix, which was derived from a compilation of 112 promoters defined by biochemical and genetic evidence (Mulligan *et al*, 1984). A consensus promoter used by *E. coli* RNA polymerase holoenzyme (σ^{70}) has a raw score of 332 and a similarity score of 100 when analysed with the same matrix.

Addendum

Since the submission of this thesis, a consensus sequence for a σ^s-selective promoter has been proposed (Becker and Hengge-Aronis, 2001). The *pcnB* promoter is similar to this consensus sequence at significant nucleotide positions. Because of this, it is believed that expression from the *pcnB* promoter may be regulated by the cellular level of the stationary phase-specific sigma factor, σ^s. The σ^s subunit of RNA polymerase is the master regulator of the general stress response in *E. coli*. It is encoded by the *rpoS* gene and, based on its molecular weight (38 kDa), is also termed σ^{38}. Although this consensus is similar to that for σ^{70}-controlled promoters, it is not identical. The nucleotide at position -13 is a C residue in more than 85% of the σ^s-dependent promoters that the consensus sequence was derived from. In addition, there seems to be a preference for either a T or G residue at the -14 position (each present in 40% of these promoters). This suggests that σ^s-containing RNA polymerase holoenzyme ($E\sigma^s$) may use an extended -10 region that includes the -13 and perhaps -14 positions. Becker and Hengge-Aronis (2001) demonstrated that a C residue at the -13 position results in the strongest expression from a σ^s-dependent promoter. The *pcnB* promoter from *E. coli* possess a C and a G at these key positions (they are also conserved in homologues of *pcnB* from other eubacteria - Section 5.2). The expression of *pcnB* might, therefore, increase as the cell enters stationary phase, when the level of σ^s increases rapidly and strongly in response to a variety of different stress conditions (such as nutritional deprivation and changes in the osmolarity). This hypothesis could be tested by measuring the β-galactosidase activity from a *pcnB'-lacZ* transcriptional fusion (such as λNB21) at hourly intervals from lag phase through to stationary phase.

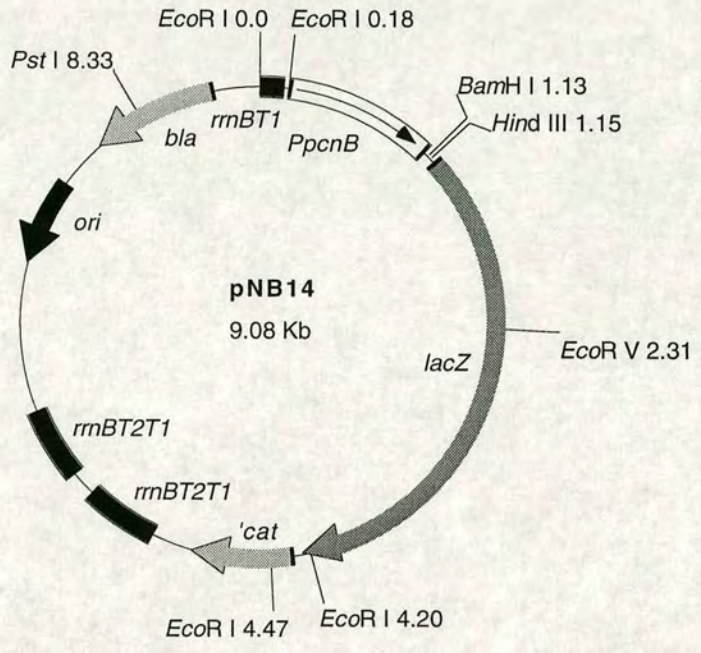
3.3 The relative strength of the *pcnB* promoter

Close resemblance to a consensus σ^{70} region suggested that the promoter might be fairly strong. However, the overall strength of a promoter cannot be predicted entirely from the base composition of the -35 and -10 regions: the sequence immediately around the start site and the initial transcribed region (from +1 to +30) are also known to influence promoter strength. A relative indicator of *pcnB* promoter strength was obtained by comparing the β-galactosidase activity from a *pcnB-lacZ* transcriptional fusion with other promoter-reporter constructs. The various fusions I made were in pASHOK, a derivative of the promoter analysis vector, pKK232-8, in which the first eight codons from the 5'-terminus of the *cat* reporter gene were deleted and replaced by *lacZ* (A. Kumar, *pers. comm.*).

3.3.1 Construction of several promoter-*lacZ* transcriptional fusions, including P_{pcnB}'-*lacZ*, in the multicopy vector, pASHOK

The 947 bp *Eco*R V fragment from pJM513 was cloned into the *Sma* I site of pASHOK. The resultant construct, pNB14 (Figure 3.3.1), forms intense blue colonies when cells harbouring it are grown on media containing X-gal. This supports the idea that a promoter is contained within the pNB14 insert, which is in the appropriate orientation to initiate *lacZ* transcription. A product of the same cloning event, pNB16 is an otherwise identical plasmid in which the *EcoR* V fragment ligated in the opposite orientation. This construct forms pale blue colonies when cells harbouring it are grown on media containing X-gal.

Figure 3.3.1 pNB14.



pNB14 is a $P_{pcnB'}$-*lacZ* transcriptional fusion construct based on pASHOK, itself a derivative of pKK232-8. See text for construction details. The presence of a single in-frame amber stop codon immediately downstream from the *Hin*d III site should prohibit translational read-through from *pcnB'* into the *lacZ* gene. This will ensure that the translation of *pcnB'* will not affect *lacZ* expression. *rrnBT1T2*, ribosomal RNA T1 and T2 transcription terminators; *bla*, ampicillin resistance gene; *'cat*, chlorampenicol resistance gene with N-terminal region deleted.

In addition to these two multi-copy *pcnB'-lacZ* transcriptional fusions, I cloned two other promoters, previously characterised in other systems, into pASHOK: pNB4 carries P_{cons}, a σ^{70} consensus promoter. This was constructed by annealing two oligonucleotides, 091E and 092E obtained from R. Hayward, (Figure 3.3.2 for sequence) and cloning the synthetic duplex into the *Bam*H I and *Hin*d III sites of pASHOK.

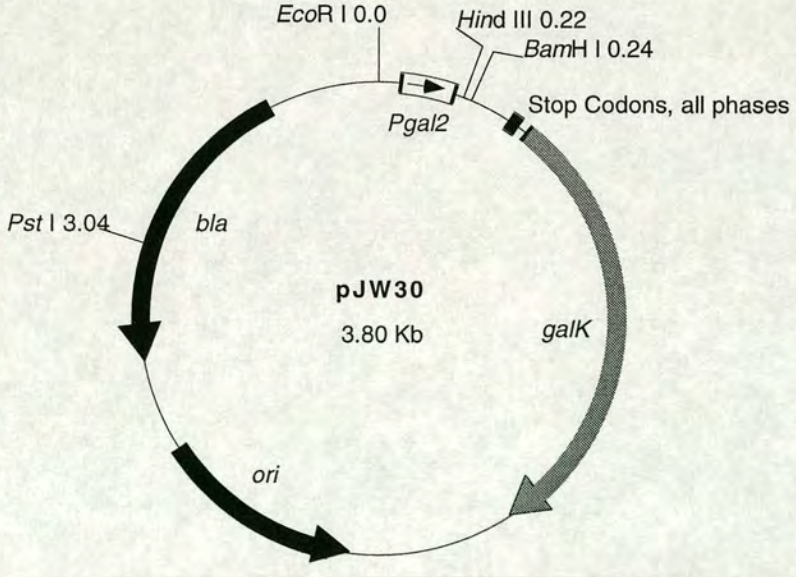
Figure 3.3.2 Nucleotide sequence of P_{cons}.

```
                          -35                    -10
092E      5' - GATCCATGTTTGACAGGTACCCGGAGCTCGAGTATAATA - 3'
                 ||||||||||||||||||||||||||||||||||||
091E           3' - GTACAAACTGTCCATGGGCCTCGAGCTCATATTATTCGA - 5'
```

P_{cons}. Duplex sequence includes several restriction sites including *Bam*H I and *Hin*d III (5'- and 3'- cohesive ends respectively). Underlined nucleotides indicate -35 and -10 σ^{70} consensus regions.

The *galE* promoter P_{gal2} (derived from mutant DNA with P_{gal1} inactivated, and the *galE* ribosome binding site deleted to prevent translation that might otherwise interfere with the *lacZ* reporter) was cut from pJW30 (Figure 3.3.3; Wright *et al*, 1992) as a 0.85 kb *Pst* I/*Hin*d III fragment and cloned as a replacement between the corresponding sites in pASHOK, creating pNB18. This plasmid was constructed in order to compare the activity of the P_{gal2} promoter in two different reporter systems (*galK* and *lacZ*) and, by using this as a standard, to allow a comparison of the *pcnB* promoter with a wide selection of other promoters previously characterised in the *galK* system.

Figure 3.3.3 pJW30.



Although the upstream transcriptional terminator present in pASHOK (between nucleotides 0 and 179) was deleted during the construction of pNB18, this does not preclude its comparison with pJW30, which also lacks this control element and is indeed identical in upstream DNA sequence from *Eco*R I to beyond *ori*.

3.3 2 The *pcnB* promoter is as efficient as a synthetic σ^{70} consensus promoter

The β-galactosidase activity from the promoter-*lacZ* fusions described in the previous section are presented in Table 3.3.1.

Table 3.3.1 Quantitative analysis of *pcnB* promoter strength compared to the promoters of several other genes.

Promoter	GalK Units pKO-based construct	β-galactosidase (MU) pASHOK-based construct[c]	β-galactosidase (MU) λRS45-based construct[d]
lacZYA	500-1000[a]	-	474 ± 13[b]
trp	ca. 2000	-	-
tac	ca. 4000	-	-
dnaA	-	-	168 ± 1.7 (λRWS945)
gal2	158 ± 3 (pJW30)	960 ± 47 (pNB18)	-
σ^{70} Consensus	-	1162 ± 28 (pNB4)	-
None	-	< 0.01 (pASHOK)	1.5 ± 0.01 (λRS551)
pcnB (5'>3')	-	1127 ± 31 (pNB14)	55 ± 0.9 (λNB21)
pcnB (3'>5')	-	2.1 ± 0.7 (pNB16)	-

Comparative analysis of promoter stengths. β-galactosidase assays were carried out according to Miller (1972) in strain P90C, harbouring *lacZ* fusions based on a multi-copy plasmid (pASHOK) or a monolysogenic phage lambda (λRS45). Values for each construct were obtained from cells growing exponentially in LB broth + ampicillin (50 μg ml^{-1}) at 37°C and are the average of three measurements ± the standard deviation. GalK data courtesy of R. Hayward. MU, Miller Units; **a**, depending on growth conditions; **b**, Data were obtained from chromosomally encoded P_{lacZYA} in MG1655; **c**, See text for details of construction. **d**, Data included here for comparison.

Based on the data in Table 3.3.1, it is estimated that the *pcnB* promoter is approximately 22 times less active than the well characterised *ptac* promoter and between 2.5 - 5 times weaker, depending upon growth conditions, than the *lacZYA* promoter. The *pcnB* promoter appears to be as strong as a synthetic consensus promoter. Although this might suggest that the *pcnB* promoter is relatively strong, this conclusion must be placed in the context of a recent report, which demonstrated that a perfect consensus may prevent the RNA polymerase from quickly clearing the promoter (after initiation), due to the high affinity interaction (Ellinger *et al*, 1994).

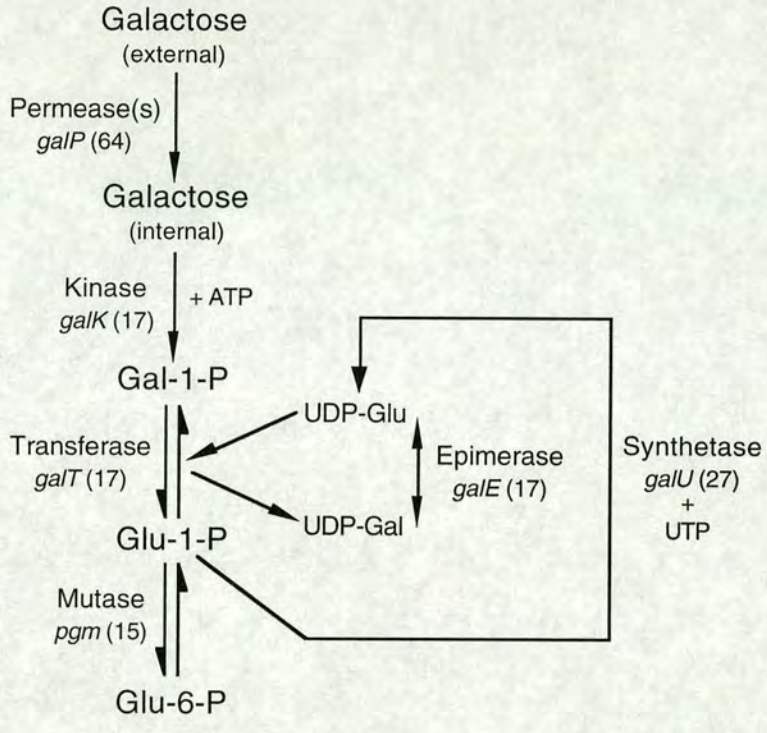
3.4 A temperature sensitive mutant of *pcnB* would aid the study of its post-transcriptional regulation

A universal feature of all reporter gene fusions is the non-native state of the 3′ end of the resulting hybrid mRNA. This is not normally a problem, but when the subject of study is an RNA processing enzyme, the results from such work may be unreliable. The most likely substrate for the autogenous control of *pcnB* expression would therefore be missing from any *pcnB′-lacZ* hybrid message. A more reliable picture of any post-transcriptional control events, based on the known activity of PcnB, would be obtained by directly comparing the stability of *pcnB* mRNA in isogenic *pcnB*$^+$ and Δ*pcnB* strains. Such an approach would require the use of a *pcnB* mutant able to synthesise full-length *pcnB* mRNA, without the encoded polypeptide being biologically active; complete translation of the message is desirable as an mRNA not coated with translating ribosomes is exposed and potentially susceptible to endo-ribonuclease degradation (Iost and Dreyfus, 1995), though this is not always the case (Lopez *et al*, 1998; Wagner *et al*, 1994). The most appropriate mutant for this study is, therefore, a conditional *pcnB* ts strain. As such a mutant was not known to be available, I devised a scheme to islotate one utilising the accumulation of toxic intermediates of galactose metabolism as a means for positive selection.

3.4.1 A strategy for the isolation of *pcnB* temperature sensitive mutants

The galactose-glucose interconversion pathway in *E. coli* (Adhya, 1987; Horecker *et al*, 1960; also know as the Leloir Pathway) identifies a number of enzymes which are involved in the uptake of galactose, and its subsequent metabolic conversion to glucose-6-phosphate. Figure 3.4.1 shows the structural genes and their cognate enzymes present in this pathway.

Figure 3.4.1 The galactose-glucose interconversion pathway in *E. coli*.



The galactose-glucose interconversion pathway in *E. coli*. Galactose is converted into glucose 6-phosphate in four steps. The first three reactions result in a net production of glucose 1-phosphate. The final step is the isomerization of glucose 1-phosphate to glucose 6-phosphate. Numbers in parenthesis are the locations of the genes, in minutes, on the *E. coli* chromosome.

A mutation in any of these genes prevents *E.coli* from growing on D-galactose as the sole carbon source. An accumulation of UDP-galactose (and galactose-1-P) due to a mutation in the *galE* gene, encoding the epimerase uridinediphosphogalactose-4-epimerase, is bactericidal. The presence of glucose in the medium prevents induction of the galactose operon by inducer exclusion, a component of the 'glucose effect' that inhibits the transport of inducers (D-galactose in this case) into the cell. Thus, when a *galE* mutant is grown on glucose-free minimal medium in the presence of D-galactose, the cells are killed by lysis. A strain with a *galK*(am), *galE* background harbouring a ColE1-related plasmid encoding an appropriate amber suppressor, should, when grown on minimal-galactose, accumulate a lysis-inducing level of UDP-galactose. The loss of the amber suppressor-

encoding plasmid (or a reduction in the copy number of such a plasmid) should lead to the complete (or partial) inactivation of *galK*, and thus survival of the cell. Such a loss or significant reduction in plasmid copy number could occur by the acquisition of a mutation within *pcnB*. Temperature-sensitive alleles of *pcnB* could be isolated by selecting for viable transductants (derived from a hydroxylamine-mutagenised P1 lysate) at 42°C on minimal-galactose, and screening for inviability at 30°C.

3.4.2 Construction of a *galK*(am), *galE* strain

The first step was to construct a strain with the desired *galK*(am), *galE* background. Initially, this was to be achieved by moving the *galK*(am), *galE* locus from OV2 by P1 transduction into MG1655. This transfer is essential in order to separate these alleles from a temperature sensitive amber suppressor mutation (*tyrT181*) in the tyrosine tRNA gene, located at 27.7 minutes on the *E. coli* chromosome. OV2 was transduced with a lysate made on strain *nadA302*::*kan*^{R}. This produced transductants which should have a *kan* gene to*galK*(am)/*galE* linkage of 75% (*nadA* is at 16.8 minutes and *galK*/*galE* are at 17 minutes on the *E. coli* chromosome). Having introduced a selectable marker, it was now possible to transfer the *kan*^{R}/*galK*(am)/*galE* linkage group to the new background. Before this was done, it was necessary to confirm that the *galK*(am)/*galE* mutation in OV2 had not been replaced by the cognate wild-type alleles from the *nadA* donor strain. The kanamycin resistant transductants were plated onto MacConkey galactose (1%) at 30°C and 42°C. For unknown reasons, the desired phenotype (no growth at 30°C, colourless growth at 42°C) was not obtained.

A second approach to constructing the desired strain was the reverse of the above procedure. Instead of moving the *galK*(am)/*galE* locus to a new background to separate it from *tyrT181*, I eliminated the *tyrT181* allele by co-transducing OV2 to Trp^{+}, TyrT^{+} with a P1 lysate made on MG1655 (*tyrT* and *trp* are located on the chromosome at 27.7 and 28.3 minutes, respectively, with a linkage of 16.6%). Successful removal of the unwanted allele was confirmed by scoring Trp^{+} transductants that were colourless at 30°C on MacConkey-arabinose (because OV2 is *ara38*(am).

3.4.3 Characterisation of MDO: a strain suitable for the isolation of *pcnB* mutants based on galactose toxicity

At this stage I stopped using my isolate in favour of MDO, a newly aquired strain from the *E. coli* Genetic Stock Centre at Yale University. This strain required no genetic alteration as it already possessed the desired *galK*(am)/*galE* genotype and did not encode an amber suppressing allele. Before proceeding to use this strain, all relevant genotypes were confirmed. The Gal$^-$ phenotype was demonstrated on minimal-galactose and minimal-galactose + glycerol. As this strain was not expected to utilise galactose, the additional carbon source was necessary to allow growth and indicate viability. Glycerol was used because it feeds into the glycolytic pathway (via the isomerisation of dihydroxyacetone phosphate to glyceraldehyde 3-phosphate) without inducing catabolite repression.

Before attempting to use MDO to isolate a *pcnB* ts mutant, the feasibility of the isolation scheme was tested empirically. Firstly, in order to identify potential *pcnB* mutants in this background, it was important to know the single-cell ampicillin resistance titre for MDO and MDOK (Δ*pcnB*). As different genetic backgrounds can have a pronounced effect on the behaviour of a particular phenotype, it could not be assumed that the titre for MDOK would be the same as the well documented titre for MM38K (Δ*pcnB*). All four strains (MDO, MDOK, MM38 and MM38K) were transformed with pBR325 and plated to give single colonies on LB over a range of ampicillin concentrations from 50 to 1000 μg ml^{-1}. The 50% lethal dose for MDOK was about 180 μg ml^{-1} ampicillin compared with 220 μg ml^{-1} for MM38K (Masters *et al*, 1993). Thus it appears that MDOK is more sensitive to ampicillin than MM38K. Since there is a large difference in ampicillin resistance between the Δ*pcnB* and *pcnB*$^+$ forms of MDO, it is a suitable background for the isolation of ts *pcnB* mutants (the 50% lethal dose for both MDO and MM38 is greater than 1000 μg ml^{-1} ampicillin).

3.4.4 Plasmid-encoded amber suppressor tRNAs charged with arginine, proline or cysteine allow recovery of Δ*pcnB* mutants on galactose

Since the isolation procedure is dependent on successful suppression of *galK*(am), I screened the ability of a number of ColE1-based plasmids encoding amber suppressors to suppress the *galK*(am) mutation in strains MDO and MDOK. These plasmids each encode a different amber suppressor tRNA capable of inserting one of seven different amino acids at the site of an amber codon (Interchange™ Amber Suppressor *in vivo* Mutagenesis System, Promega; Cat.# Q5080). Table 3.4.1 presents the results of this screen.

Galactose was found to be cytotoxic to strain MDO (*pcnB*$^+$)when transformed with any of the amber suppressor plasmids (Table 3.4.1). Although MDO (*pcnB*$^+$) with pCys does form colonies on minimal-galactose/glycerol, they are very small, indicating that the cells are stressed. This galactose-associated cytotoxicity appears to decrease in MDOK (Δ*pcnB*) when harbouring pArg, pPro and pCys, but not pGly, pPhe, pGlu, and pHis (Table 3.4.1). Only pCys allows 100% survival of MDOK on galactose, and is therefore likely to enhance the chances of recovery of *pcnB* mutant isolates. Presumably galactose is fully cytotoxic to MDOK when transformed with pGly, pPhe, pGlu or pHis either because the amber suppressors they encode are more efficient than pArg, pPro or pCys at allowing translation of *galK42*(am), or because the activity of the *galK42*(am) gene product is greater with the amino acids they insert. As pArg, pPro and pCys were the only suppressors that allowed growth of MDOK on minimal-galactose/glycerol, strain MDO was freshly transformed with these plasmids and subjected to random mutagenesis in order to isolate *pcnB* ts mutants.

3.4.5 Localised P1-mutagenesis of MDO (pCys) yielded a single galactose tolerant transductant

Random mutations were introduced into the 3.4 minute region by localised P1-mutagenesis (Hong and Ames, 1971; Section 2.4.7 for details). The DNA of phage P1 was randomly mutated by exposure to hydroxylamine and transferred to the chromosome by transduction. This technique allows the isolation of ts and other types of mutations in any specific small region (about 1%) of the bacterial chromosome by mutating the transducing phage DNA rather than the bacterial DNA. MDO harbouring either pArg, pPro, or

Table 3.4.1 The restoration of *galK*(am) function in strains MDO and MDOK by several different plasmid-encoded tRNA amber suppressors.

The restoration of *galK*(am) function in strains MDO and MDOK is indicated by the inability to grow on minimal-galactose/glycerol at 42°C whilst harbouring different plasmid-encoded tRNA amber suppressors. Fresh transformants of MDO and MDOK were grown exponentially in minimal-glycerol-Amp (50 μg ml^{-1}) to an OD_{540} of 0.3, diluted in fresh minimal-glycerol so as to obtain approximately 200 colonies per plate on galactose-free medium and spread on ampicillin-free minimal-glycerol and minimal-galactose/glycerol plates. The number of colonies on plates without galactose (data not shown) was defined as 100 (i.e. minimal-glycerol is a '100% growth' control); the figures in each horizontal row are percentages of these initial values. Standard three letter code indicates incorporated amino acid. pBR325 is a non-suppressor control. Galactose was used at 1%. Glycerol was used at 0.2%.

Table 3.4.1 The restoration of *galK*(am) function in strains MDO and MDOK by several different plasmid-encoded tRNA amber suppressors.

Plasmid	Strain	Number of colonies on Minimal-Galactose+ Glycerol
pArg	MDO	0
	MDOK	24
pGly	MDO	0
	MDOK	0
pPhe	MDO	0
	MDOK	0
pPro	MDO	0
	MDOK	19
pGlu	MDO	0
	MDOK	0
pCys	MDO	Very tiny; many
	MDOK	100
pHis	MDO	0
	MDOK	0
pBR325	MDO	100
	MDOK	100

pCys were transduced with a hydroxylamine-mutagenised P1 lysate made on MRi84 (*zad*::Tn*10* *pcnB*$^+$). Transductants were plated onto ampicillin-free minimal medium containing 1% galactose, 0.2% glycerol and tetracycline (10 μg ml^{-1}). The viability of a P1 lysate diminishes more quickly if it has been hydroxylamine-mutagenised. The following controls (harbouring no plasmids) were set up in parallel at 42°C to check the transducing ability of the mutagenised lysate (same volumes of lysate as actual screen):

MDO on MM-galactose/glycerol (no antibiotics or lysate): Confluent growth.

MDO on MM-galactose/glycerol + tet (no lysate): No growth.

MDO + mutagenised P1 lysate on MM-galactose/glycerol + tet: ~200 transductants.

After overnight incubation at 42°C, MDO pCys yielded one small transductant. None were obtained with pArg or pPro. The single transductant isolated on minimal-galactose/glycerol was considered a likely plasmid copy number mutant, when compared with the control yield. The isolate, designated PAP40, (which remained ampicillin resistant) was cured of pCys. A single colony was inoculated into LB broth without selection and grown overnight to stationary phase. The cells were diluted in fresh LB broth so as to obtain approximately 200 colonies per plate on ampicillin-free medium. These were then patched onto LB and LB-Amp (50 μg ml^{-1}) to look for plasmid loss.

3.4.6 Acquisition of the *pcn-40* allele (from PAP40) results in a temperature-independent reduction in plasmid copy number and RNA I half-life

A plasmid-free segregant of PAP40 was transformed with pBR325 and plated onto LB-Amp (50 μg ml^{-1}) and LB-Amp (500 μg ml^{-1}) both at 30° and 42°C. The strain grew on low ampicillin at both temperatures, whilst failing to grow on high ampicillin at 42°C. This is the expected phenotype of a *pcnB* mutant. However, the isolate did not form colonies on LB-Amp (500 μg ml^{-1}) at 30°C, which suggested that the strain had acquired a simple unconditional point mutation (designated *pcn-40*), rather than a temperature-sensitive one. Strain MM38 was transduced to *pcn-40* with a P1 lysate made on PAP40. This demonstrated that the phenotype associated with *pcn-40* is an inheritable characteristic not associated with an extra-chromosomal element. The ampicillin concentration at which the viability of MM38 (*zad*::Tn*10* *pcn-40*)

equals 50% was shown to be about 250 μg ml^{-1}. The rate of decay of RNA I in MM38 (*pcn-40*) - as shown by its half-life - was kindly determined by Dr. Uta Binnie in this laboratory and compared with the wild-type value in MM38 (*pcnB*$^+$). The two strains, both harbouring pBR325, were grown in LB broth + Amp (50 μg ml^{-1}) at 37°C to OD_{600} of approximately 0.5. Cells were treated with rifampicin, an antibiotic that inhibits new RNA synthesis, and aliquots of the cultures withdrawn for quantitative Northern analysis at the times indicated (Figure 3.4.2).

The intensity of the radioactive hybridisation signal from each band was quantified using a Molecular Dynamics PhosphorImager 400S (Molecular Dynamics). The ratio between the RNA I_{total} signal and the $tRNA^{Ser}$ signal at each time point was normalized by setting the RNA $I_{total}/tRNA^{Ser}$ ratio at time 0 to equal 1. The normalized ratios of RNA $I_{total}/tRNA^{Ser}$ were plotted as a function of time after rifampicin addition (Figure 3.4.3). Extrapolation of the fitted curves to the abscissa gives RNA I half-lives in MM38 (*pcn-40*) and MM38 (*pcnB*$^+$) of approximately 15 and four minutes, respectively. The half-life of RNA I in *pcn-40* appears to be about four-fold greater than in the wild-type strain. This demonstrates that *pcn-40* can confer the phenotype of increased RNA I stability on a strain. These data, together with the ampicillin titre, show that *pcn-40* is phenotypically similar to other *pcnB* point mutants. Its isolation demonstrates the validity of the positive selection scheme as a method for isolating *pcnB* mutants, but emphasises the need for screening to identify ts mutants among them.

Figure 3.4.2 The decay of RNA I in MM38 (*pcn-40*).

Measurement of the rate of decay of RNA I in MM38 (*pcn-40*) compared with the wild-type rate in MM38 (*pcnB*$^+$). The two strains, both harbouring pBR325, were grown in LB broth + Amp at 37°C to OD_{600} of approximately 0.5. Cells were treated with rifampicin (250 μg ml^{-1}) and aliquots of the cultures withdrawn for quantitative Northern analysis at the times indicated. The bands corresponding to RNA I_{108} and RNA I_{103}, respectively, are indicated; $tRNA^{Ser}$ is a stable RNA species that served as a convenient measure of total relative RNA concentration and was used to correct for unequal loading of the gel lanes. The filter was first probed for RNA I (using RNA II labelled with [^{32}P]-UTP during *in vitro* transcription from a T7 promoter), then for the stable $tRNA^{Ser}$ (with oligonucleotide SS-2: 5′-d[CCGGTAGAGTTGCCCCTACTCCGGTTTTAG]-3′ end labelled with [γ-^{32}P]-ATP by T4 polynucleotide kinase. Time (min), refers to time after the addition of rifampicin.

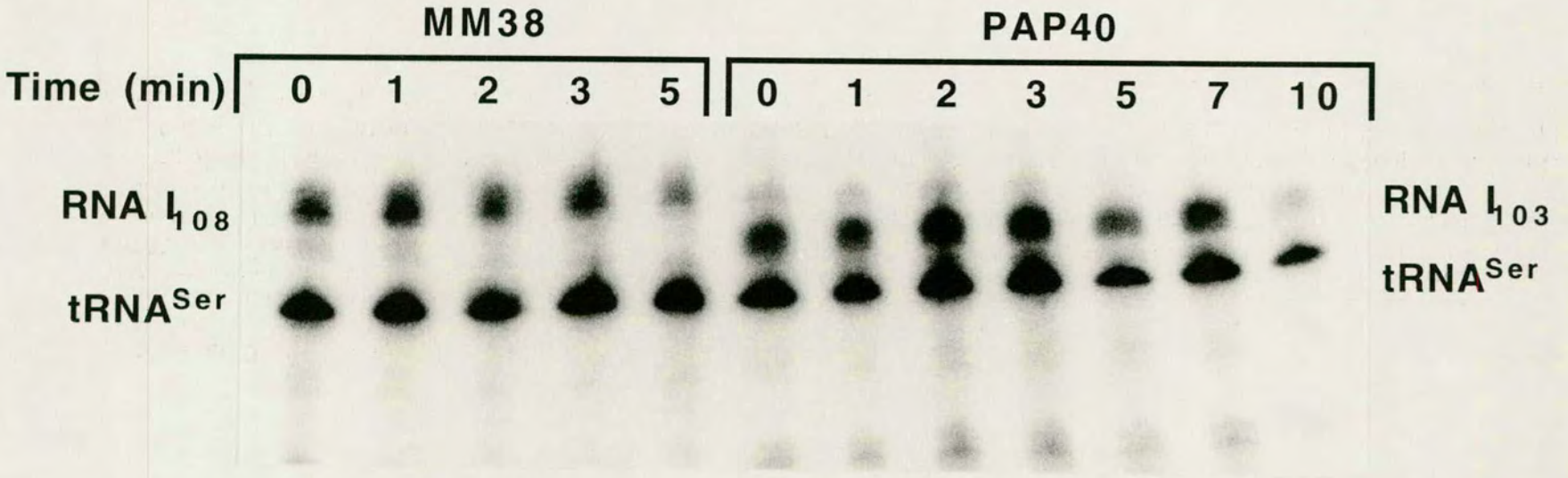
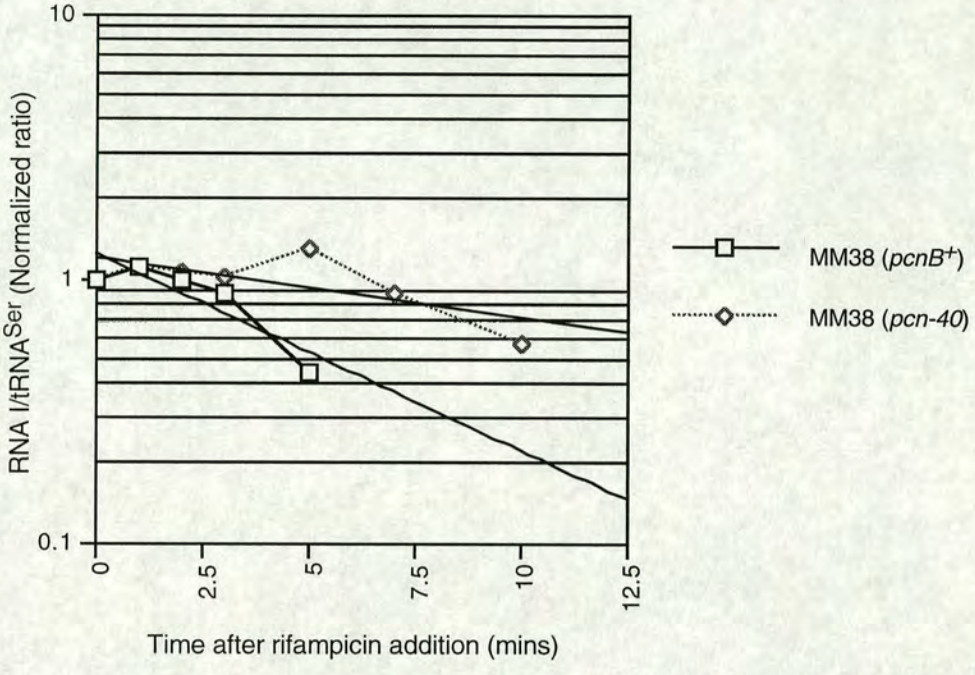

MM38
PAP40
Time (min)
0
1
2
3
5
0
1
2
3
5
7
10
RNA I108
tRNASer
RNA I103
tRNASer

Figure 3.4.3 Half-life determination of RNA I in MM38 (*pcn-40*).



Half-life determination of RNA I in MM38 (*pcn40*). Data were derived from the autoradiogram shown in Figure 3.4.2. See text for details.

Shortly after the isolation of PAP40, this ts mutant search was abandoned because attempts to identify *pcnB* mRNA produced from the chromosome (i.e. in single copy) proved unsuccessful. In addition, results (obtained after the isolation of PAP40) from strains monolysogenic for a lambda phage bearing a P_{pcnB}-*lacZYA* transcriptional fusion (described in Section 3.5.3) suggest that any regulation of the expression of *pcnB* at the level of transcription is very slight, and due to derepression of the *pcnB* promoter rather than an increase in transcript stability. Retrospectively, a less involved approach would have been to clone an unconditional point mutant, such as *pcnB80* (Lopilato *et al*, 1986), into a PcnB-insensitive multicopy plasmid (such as the pSC101-derived cloning vector, pGB2; Churchward *et al*, 1984) and compare the half-life of this mutant *pcnB* mRNA in Δ*pcnB* and *pcnB*$^+$strains. Even if there is some residual activity from the point-mutant, provided the

difference from wild-type is great enough, using this approach it should be possible to determine if *pcnB* expression is autogenously controlled by regulating its mRNA turnover. The use of an unconditional point mutant, rather than a ts mutant, would also avoid any interference that may arise from induction of the heat-shock response. Substituting the point mutant with an engineered *pcnB* amber mutant might be a viable alternative, were it not that the absence of translating ribosomes may distort the final result; the same problem would arise with an engineered frame-shift mutant that caused the introduction of stop codons. An unconditional point-mutant, borne on a multicopy vector, could also facilitate determining the size of the native *pcnB* transcript, by increasing the gene dosage of *pcnB* whilst possibly reducing the turnover of its transcript. This could also reveal whether downstream *folK* is indeed part of the same transcription unit as *pcnB*.

3.5 The effect of PcnB on the expression of its gene at the level of transcription

For the purpose of estimating the relative strength of the *pcnB* promoter, the multi-copy state of pASHOK was largely unimportant. However, the study of PcnB's role, if any, in determining the level of its own expression must avoid vectors like pASHOK, whose copy number is known to be controlled by PcnB. For this reason, I employed a single-copy phage λ-based *lacZ* transcriptional fusion vector.

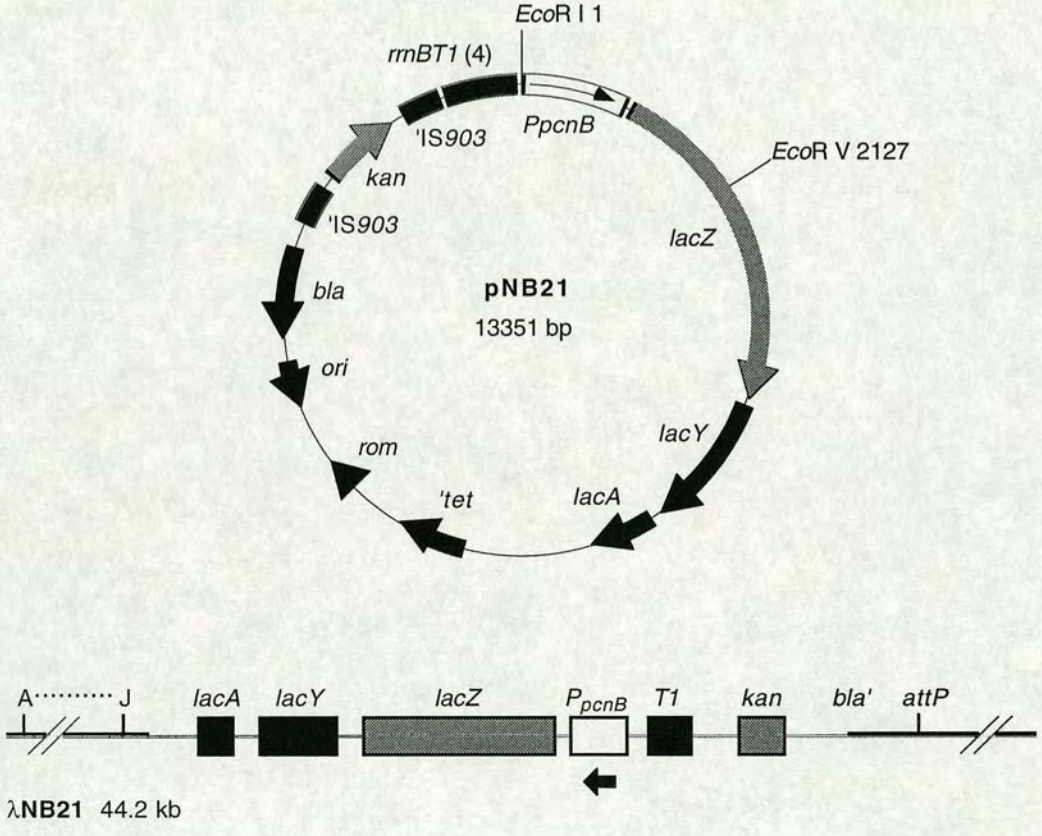
3.5.1 The vectors pRS551 and λRS45: a two component system for the cloning of *lacZ*-based transcriptional fusions in single-copy

Described by Simons *et al*, (1987), this two component system consists of a phage, λRS45, and a plasmid, pRS551. λRS45 contains *lacY* and *lacA* downstream of *'lacZ*, a large deletion from the 5'-end having left only the distal third of *lacZ* intact. The phage also contains the 5'-half of the β-lactamase gene (*bla*). pRS551 contains *lacY* and *lacA* as well as the complete *lacZ* gene. Upstream of *lacZ* there are four tandem copies of the *rrnB* T1 transcriptional terminator and a limited number of unique restriction sites in which to insert the desired promoter sequence. On the other side of the terminators there is a copy of *aphA* from Tn*903*, which confers kanamycin resistance, and a complete copy of the *bla* gene, which confers ampicillin resistance. Hence, pRS551 carries homology to λRS45 on either side of the cloning site and a double recombination event between these homologous sequences will yield a phage carrying any promoter sequence which was inserted into pRS551 transcriptionally fused to *lacZ*. The presence of four strong transcriptional terminators immediately upstream of the cloned promoter eliminates the possibility that transcripts originating elsewhere, either in the vector or on the chromosome, can interfere with or contribute to *lacZ* expression. Additionally, the recombinant phage will carry *aphA* conferring kanamycin resistance, a property which facilitates selection of lysogens of the recombinant phage. It will, however, not confer ampicillin resistance as the recombination event will not result in a complete copy of the *bla* gene on the phage.

3.5.2 Construction of λNB21, a P_{pcnB}'-*lacZ* transcriptional fusion derived from pRS551 and λRS45

The 2.1 kb *Eco*R I/*Eco*R V fragment containing the *pcnB* promoter region was cut from pNB14 (Figure 3.3.1) and cloned between the corresponding unique sites in pRS551. The resultant plasmid containing the *pcnB* promoter transcribing towards *lacZ* was named pNB21 (Figure 3.5.1). Wild-type β-galactosidase is translated from its own translation initiation site in this fusion, so only regulatory processes that affect transcription from the *pcnB* promoter will change the expression of β-galactosidase from pNB21.

Figure 3.5.1 pNB21 and λNB21.



pNB21 and λNB21. Based on pRS551 and λRS45. T1, four tandem copies of the strong *rrnB* T1 transcriptional terminator, which eliminate the possibility that transcripts originating elsewhere, either in the vector or on the chromosome, interfere with or contribute to *lacZ* expression. See text for construction details.

The presence of a single in-frame amber stop codon immediately upstream from *lacZ* should prohibit translational read-through from *pcnB'* into the *lacZ* gene. This will ensure that the translation of *pcnB'* will not affect *lacZ* expression. This stop codon element was acquired from the 2.1 kb *Eco*R I/*Eco*R V fragment and originated from pASHOK. pNB21 was transformed into NM621/λRS45 which carries λRS45 as a prophage, to allow *in vivo* recombination between the plasmid and the phage. Recombinants were selected by inducing the transformed strain with UV light and infecting P90C with the resulting lysate as described in Section 2.4.3. Lysogens of the recombinant phage gave rise to Kan^R colonies. As the lysate prepared from NM621/λRS45(pNB21) contains predominantly unrecombined phage there is a high probability that many of the Kan^R lysogens will contain λRS45 in addition to recombinant phage. Therefore a new lysate was prepared by induction of a Kan^R lysogen and used to infect P90C at a multiplicitiy of infection between 10^{-6} and 10^{-2}. Phage DNA was prepared from this lysate and amplified by PCR, in order to confirm that the 2.1 kb *Eco*R I/*Eco*R V fragment from pNB21 had recombined with λRS45. Kan^R lysogens from this second (purification) round of infection were selected and eight colonies were taken from the plate containing lysogens from the infection at the lowest multiplicity of input phage. Lysogens harbouring a single copy (as opposed to multiple copies) of the recombinant prophage were initially identified by determining their β-galactosidase activity. Those lysogens with the lowest activity were considered to have a single copy of the integrated prophage. The single-copy status of these lysogens was confirmed by PCR (Powell *et al*, 1994). One of them was selected for further experiments, its prophage being named λNB21 (Figure 3.5.1). An equivalent negative control, λRS551, was created by simply transforming parental pRS551 into NM621/λRS45 and selecting for Kan^R recombinants as before.

3.5.3 Expression of β-galactosidase from the *P_{pcnB}'-lacZ* transcriptional fusion in λNB21 is only slightly derepressed in a Δ*pcnB* background

Much of the data accumulated on *pcnB* in this laboratory was obtained from work on strain MM38 (Masters *et al*, 1993). To facilitate comparisons, this strain was chosen as the genetic background for studies to establish whether *pcnB* expression is subject to autogenous control at the level of transcription. The β-galactosidase activity from λNB21 in a *pcnB* strain was compared with its activity in an isogenic wild-type background. For a meaningful comparison, it was essential to ensure that the prophage status of both strains was identical. This was achieved by initially infecting MM38 (*pcnB*$^+$) with λNB21. The resultant lysogen was transduced to Δ*pcnB* with a P1 lysate made on MM44, forming MM38KT. Based on strain MRi84 (Lopilato *et al*, 1986), MM44 has two relevant characteristics: a Tn*10* insertion showing 90% linkage to *pcnB* by P1 transduction (*zad*::Tn*10*) and a deletion of *pcnB* and its replacement with the Kan^R cassette (by a P1 lysate made on IR8903; Masters *et al*, 1993). The Tet resistance gene encoded by the Tn*10* was necessary to allow selection of Δ*pcnB* transductants in the lysogens harbouring kanamycin resistant phage λ. The tetracycline resistant transductants were confirmed to be *pcnB*$^-$ by transforming them with pBR325 and measuring the level of single-cell ampicillin resistance. Other lysogens were similarly created by infecting MM38 with λRS551 and λRWS945. Because *pcnB* has a role in determining plasmid copy number, it was considered possible that the presence of a *pcnB*-sensitive plasmid may influence the level of expression of *pcnB* by some feedback mechanism. The results of these studies are presented in Table 3.5.1.

Table 3.5.1 β-galactosidase activity of λNB21 in isogenic *pcnB*$^+$ and Δ*pcnB* strains.

Promoter	Strain	*pcnB* genotype	β-galactosidase Activity (MU)
pcnB	MM38/λNB21	+	55.7 ± 0.9
pcnB	MM38KT/λNB21	deletion	62.5 ± 1.2
pcnB	MM38/λNB21/pBR325	+	55.8 ± 1.2
pcnB	MM38KT/λNB21/pBR325	deletion	63.1 ± 1.4
lacZ	MG1655*	+	474.4 ± 13.3
lacZ	MG1655K*	deletion	479.2 ± 11.5
dnaA	MM38/λRWS945	+	168.5 ± 1.9
dnaA	MM38KT/λRWS945	deletion	165.2 ± 2.1
None	MM38/λRS551	+	1.6 ± 0.01
None	MM38KT/λRS551	deletion	1.4 ± 0.01

β-galactosidase assays were carried out according to Miller (1972) in the indicated strains, harbouring *lacZ* fusions based on a monolysogenic phage lambda (λRS45). Values for each construct were obtained from cells growing exponentially in LB broth + 0.2% glucose (and, where appropriate, kanamycin and/or ampicillin at 50 μg ml^{-1}) at 37°C and are the averages of three measurements ± the standard deviation. *, Data were obtained from chromosomally encoded P_{lacZYA}; MU, Miller Units.

The β-galactosidase activity from P_{pcnB}-*lacZYA* (λNB21) increased by a modest 11% in the absence of PcnB (Table 3.5.1). Given the ability of PcnB to promote RNA I decay by polyadenylation at its 3' end (He *et al*, 1993; Xu *et al*, 1993), it might be expected that any possible autogenous regulation of *pcnB* expression could involve this activity, by which the absence of *pcnB* results in an increase in the stability of the *pcnB* message and consequently increased levels of *pcnB*. If the increase in *lacZ* activity in the Δ*pcnB* strain is due to the absence of PcnB-mediated polyadenylation (and hence increased message stability) of the P_{pcnB}-*lacZYA* mRNA, then the same effect should be observed when examining the expression of other, non-related, *lacZYA* transcriptional fusions. However, the level of β-galactosidase activity from a

P_{dnaA}-*lacZYA* fusion was the same in a *pcnB*-null mutant and *pcnB*+ strain. It is likely, therefore, that the 11% elevated β-galactosidase activity observed in the *pcnB* null mutant is due to derepression of the *pcnB* promoter, rather than an increase in transcript stability.

The data in Table 3.5.1 also shows that the presence of a *pcnB*-sensitive plasmid has no effect on the level of transcriptional activity from the *pcnB* promoter. It should be noted that these experiments are not investigating possible autogenous control at the translational level, as the presence of an in-frame amber stop codon between the cloned fragment containing P_{pcnB} and *lacZ* should block any translational read-through into the reporter gene.

3.6 Summary

The *pcnB* promoter has been identified and characterised. It was approximately located by observing the phenotypic differences between two plasmid constructs encoding different amounts of untranslated *pcnB* sequence. Specifically, one construct seemed to lack the -35 region of an otherwise near-consensus σ^{70} promoter. By exploiting a fortuitously located restriction site, a derivative construct that spanned the -35 region with a minimum of additional upstream sequence was shown to possess the same phenotype as the full-length parental plasmid. Primer extension analysis was used to accurately map the location of the promoter. The relative strength of the *pcnB* promoter was determined by constructing a plasmid encoding a P_{pcnB}-*lacZ* transcriptional fusion, and comparing its activity to a number of heterologous promoter-*lacZ* fusions cloned in the same vector; it was found to be as efficient as P_{gal2} or a synthetic σ^{70} promoter. To facilitate a study of the post-transcriptional regulation of *pcnB*, an attempt was made to isolate a temperature-sensitive allele. The bactericidal phenotype of a *galK* (am), *galE* strain grown on galactose was rescued by the presence of an amber suppressor-encoding plasmid. Transductants (infected with a mutagenised P1 lysate) of this strain were positively selected at 42°C by their ability to survive on galactose, implying a reduction in the suppressor-encoding plasmid copy number. A mutant allele, *pcn-40*, was isolated by this method but found to be temperature-independent. In order to demonstrate whether *pcnB* expression is subject to autogenous control at the level of transcription, the β-galactosidase activity from a P_{pcnB}-*lacZYA* fusion (constructed in a lambda phage and maintained as a monolysogen to avoid variations in copy number) was measured in Δ*pcnB* and *pcnB*$^+$ backgrounds. A slight (11%) increase in β-galactosidase activity was observed from this fusion in the absence of PcnB. When compared to the data obtained from the promoter of an unrelated gene (*dnaA*) fused to *lacZYA* (and therefore encoding the same 3′ end), this result suggests that the elevated β-galactosidase activity observed in the Δ*pcnB* background is due to mild derepression of the *pcnB* promoter, rather than an increase in transcript stability.

CHAPTER 4

E. coli poly(A) polymerase I: a novel translation initiation region

4.1 Introduction

Translation when regulated is, like transcription, often controlled at the stage of initiation. Preventing unnecessary polypeptide synthesis is clearly advantageous to an organism as translation is a highly energetic process (Jacques and Dreyfus, 1990): four high-energy phosphate bonds (in the form of GTP and ATP) are hydrolysed for each amino acid that is incorporated into a growing polypeptide chain. This amounts to a total phosphate group transfer potential of $\Delta G^{\circ\prime}$ -29.2 kcal mol^{-1} (4 x -7.3 kcal mol^{-1}). For example, to synthesise a protein of 100 amino acids:

100 aminoacyl-tRNAs (charged)	uses 2 ~P each from ATP	= 200
1 initiation complex	uses 1 ~P from GTP	= 1
99 tRNAs binding to the A site	uses 1 ~P each from GTP	= 99
99 peptide bonds	uses 0 ~P	= 0
99 translocations (A to P)	uses 1 ~P each from GTP	= 99
1 termination (hydrolysis)	uses 1 ~P from GTP	= 1
	Total ~P per 100 amino acids	= 400

Having found no clear-cut evidence that the expression of *pcnB* is autogenously controlled at the level of transcription, I investigated the possibility that *pcnB* might autoregulate its expression at the level of translation. Typically, this could occur if PcnB were to bind to a region on its mRNA that overlaps the ribosome binding site, and occlude its initiation codon. Most known examples of this type of regulation encode proteins which possess, as part of their intrinsic biological function, a nucleic acid binding activity. The ability of the protein to control its translation is therefore thought to depend upon its binding to a region on its own mRNA that shares some similarity with its primary nucleic acid substrate. PcnB itself has been shown to possess poly(A) polymerase activity (Cao and Sarkar, 1992) and its N-terminus shares strong local similarity with the N-terminus of tRNA nucleotidyltransferase, another nucleic acid binding protein (Masters *et al* 1990). The substrates of these proteins include small RNAs with significant secondary structure. It is therefore conceivable that PcnB might regulate its expression by binding to a stem-loop structure in its own mRNA.

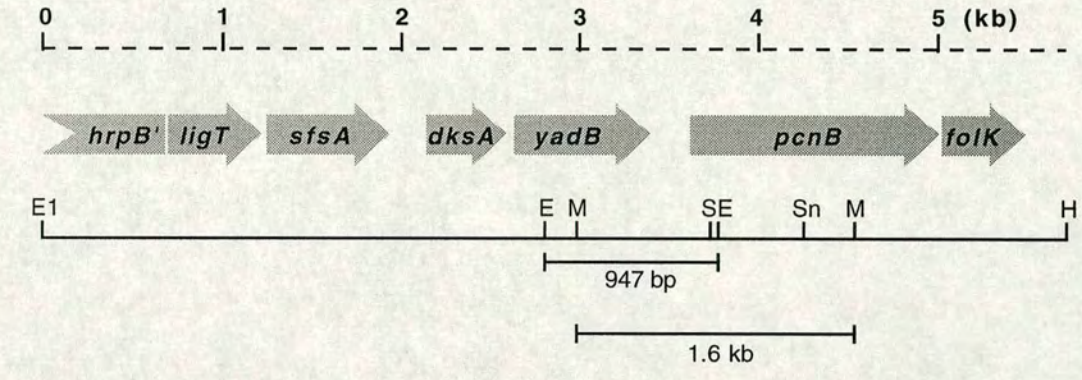
4.2 The effect of PcnB on the expression of its gene at the level of translation

To investigate the possibility that *pcnB* expression is regulated by autogenous control, the β-galactosidase activity from a lambda-based *pcnB'-'lacZ* translational fusion in a *pcnB* background was compared with its activity in an isogenic wild-type strain. The *lac* reporter system employed was essentially the same one described in section 3.5.1. In this instance, the transcriptional fusion vector used previously was substituted by pRS552, a translational fusion vector.

4.2.1 Construction of λNB25, a *pcnB'-'lacZ* translational fusion derived from pRS552 and λRS45

The translational fusion vector, pRS552, is identical to pRS551 except for a 120 bp deletion, which removes the first eight codons of *lacZ* and the intervening sequence between it and the restricted cloning site. I cloned the 947 bp *Eco*R V fragment from pJM513 (Figure 4.2.1) into the unique *Bam*H I site (end-filled with the Klenow fragment of DNA Polymerase I) in pRS552, forming pNB25 (Figure 4.2.2). This created an in-frame fusion of the first 43 codons from *pcnB* to the ninth codon of *lacZ* and was confirmed by sequencing the junction site (Figure 4.2.3). The hybrid PcnB'-'β-galactosidase is expressed from the native *pcnB* promoter and translation initiation site. Unlike the transcriptional fusion carried by λNB21, the translational fusion pNB25 should be sensitive to regulatory processes that affect both the transcription and translation of *pcnB*. The pNB25 fusion was recombined *in vivo* with λRS45, as described in section 3.5.2, to form λNB25 (Figure 4.2.2).

Figure 4.2.1 pJM513 Insert.



The pJM513 insert indicating cloned fragments referred to in this section. See text for details. Relevant restriction sites: E1, *Eco*R I; E, *Eco*R V; H, *Hin*d III; M, *Mlu* I; Sn, *Sna*B I; S, *Sph* I.

Figure 4.2.3 Fusion site at the junction of *pcnB'-'lacZ* in pNB25.

Dideoxy-sequence ladder (from double-stranded template) confirming the presence of the in-frame fusion created at the site of the *Eco*R V-*Bam*H I junction between codon 43 of *pcnB* and codon 9 of *lacZ* in pNB25. See text for construction details.

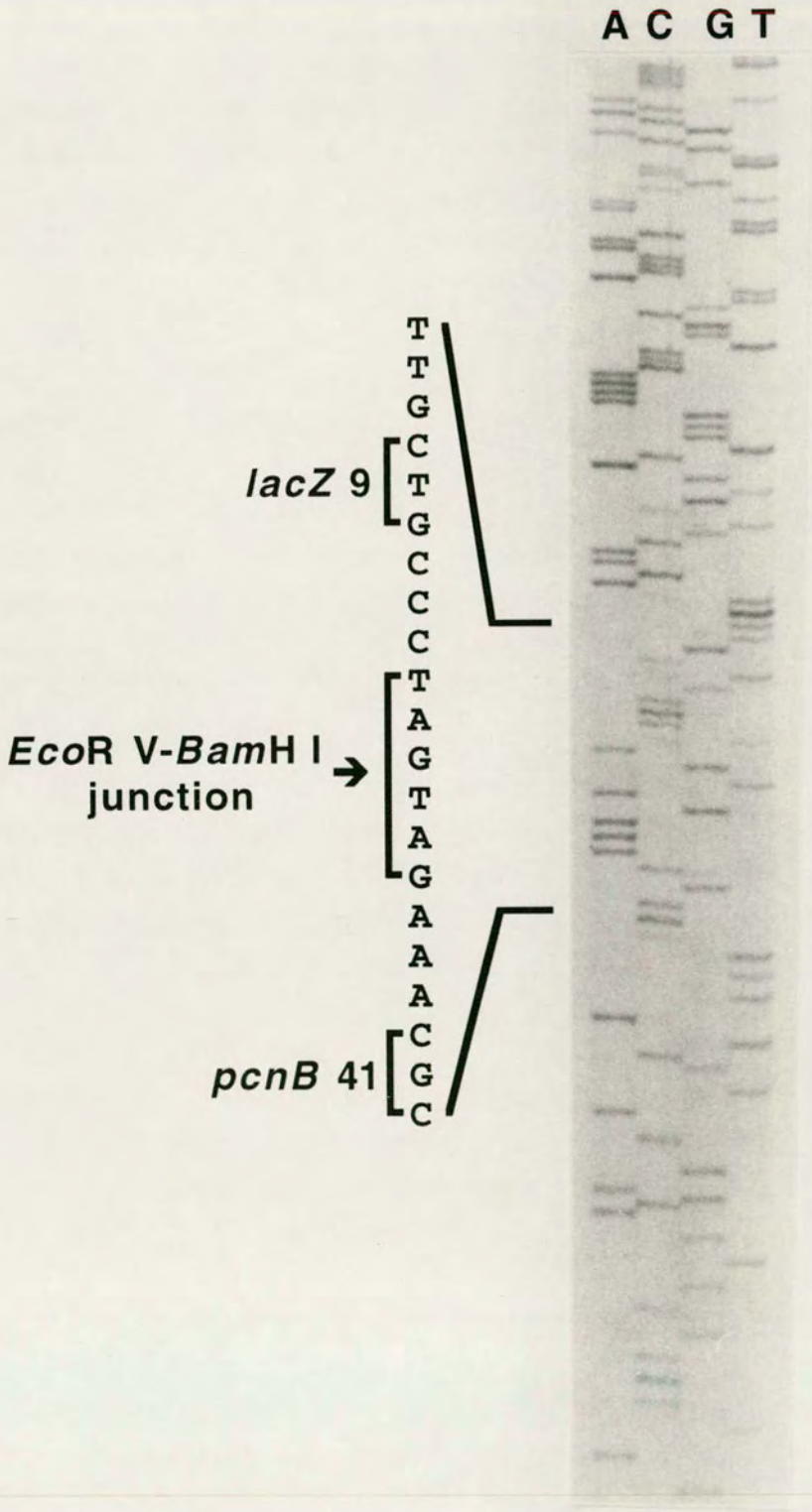
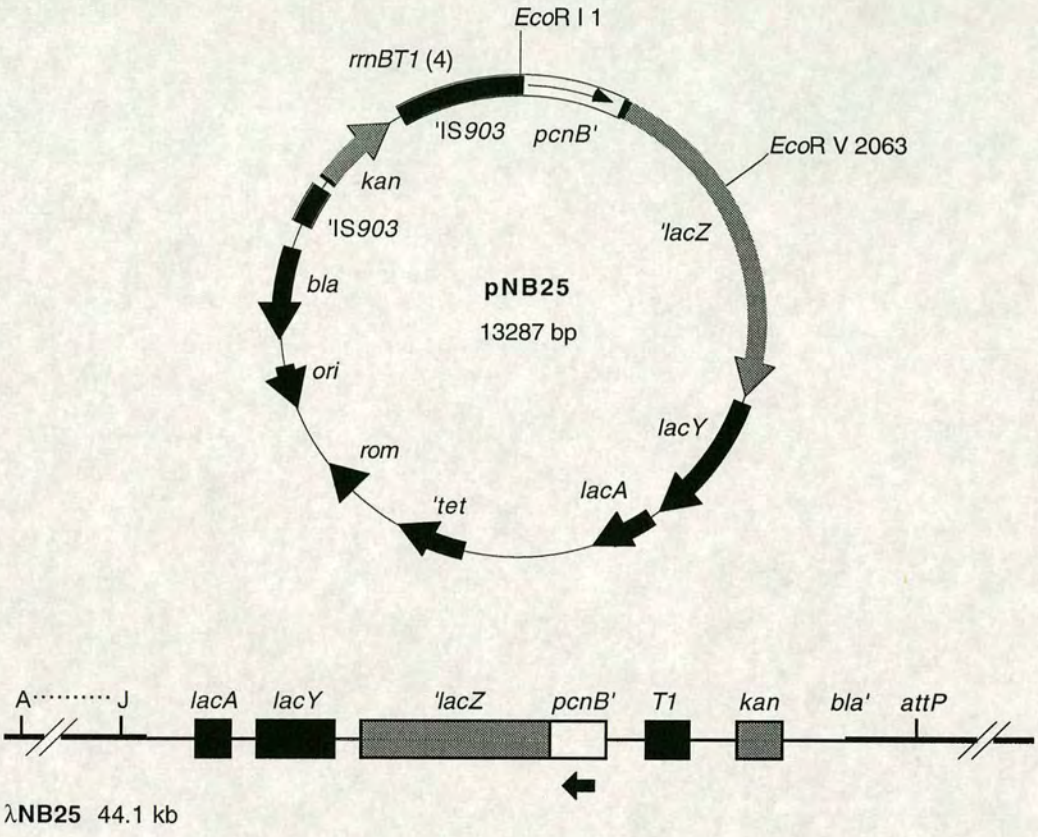

A C G T
T T G C T G C C C C T A G T A G G A A A C G C
lacZ 9
EcoR V-BamH I junction
pcnB 41

Figure 4.2.2 pNB25 and λNB25.



pNB25 and λNB25. Based, respectively, on pRS552 and λRS45. See text for construction details.

To facilitate meaningful comparisons with other studies in this lab and earlier data from this work, MM38 was again chosen as the genetic background for studies to measure the β-galactosidase activity expressed from λNB25. This strain was infected with λNB25 and transduced to Δ*pcnB* with a P1 lysate made on MM44, forming MM38KT/λNB25 (section 3.5.3). Lysogenisation prior to transduction ensures that the prophage status of MM38/λNB25 and MM38KT/λNB25 is identical.

4.2.2 Expression of β-galactosidase from the *pcnB'-'lacZ* translational fusion in λNB25 is unaltered in a Δ*pcnB* background

The absence of functional PcnB protein does not appear to have any effect on the expression of *pcnB*, as indicated by the level of β-galactosidase activity detected in the lysogens of λNB25 (Table 4.2.1). Likewise, the level of expression of *pcnB* does not appear to be influenced by a cellular excess of its gene product (originating from pJM513). Together, these results indicate that *pcnB* expression is not autogenously regulated at the level of translation. The presence of a *pcnB*-sensitive plasmid (pBR325), also has no effect on the expression of *pcnB* at the level of translation.

Table 4.2.1 The effect of PcnB on the β-galactosidase activity of λNB25, a *pcnB'-'lacZ* translational fusion vector integrated in the host genome in single-copy.

Strain	*pcnB* genotype	β-galactosidase activity (MU)
MM38/λNB25	+	82 ± 1.5
MM38KT//λNB25	deletion	81 ± 1.2
MM38/λNB25/pBR325	+	78 ± 0.9
MM38KT/λNB25/pBR325	deletion	75 ± 0.8
MM38λNB25/pJM513	+++	86 ± 1.6
MM38//λRS552	+	0.05
MM38KT/λRS552	deletion	0.05

β-galactosidase assays were carried out according to Miller (1972) in the indicated strains, harbouring *lacZ* fusions based on a monolysogenic phage lambda (λRS45). Values (in Miller Units) for each construct were obtained from cells growing exponentially in LB broth + 0.2% glucose (and, where appropriate, ampicillin at 50 μg ml^{-1}) at 37°C and are the averages of three measurements ± the standard deviation. The background level of β-galactosidase activity was determined in lysogens of λRS552 (the insertless parental vector, pRS552, transferred to λRS45).

4.3 Identification of the *pcnB* translational initiation codon

It has been postulated that the translational start site of *pcnB* is the unusual UUG codon. (Cao and Sarkar, 1992). A systematic study of the effect of changes to an initiation codon from UUG to GUG or AUG on the translational efficiency of *cya* revealed that the native UUG codon was responsible for the lowest efficiency of translation (Reddy *et al*, 1985). It was concluded that the occurrence of the UUG initiation codon provides, at least in this case, a mechanism for limiting the expression of *cya*. This conclusion is consistent with the finding that the cellular concentration of adenylate cyclase is very low. Is this also a factor in determining the level of *pcnB* expression? Difficulties with identifying PcnB *in vivo* have certainly been encountered in this laboratory, suggesting it is expressed at low levels.

Initially, I set out to confirm whether the translation of *pcnB* was initiated from the proposed UUG codon. The method involved introducing amber stop codons by site-directed mutagenesis at selected sites within and outside the putative open reading frame. Plasmids encoding these different *pcnB*(am) mutants were then subjected to *in vitro* coupled transcription/translation. An absence of full-length PcnB would indicate that the amber mutation lay within the open reading frame. In this way, the initiation codon was systematically mapped.

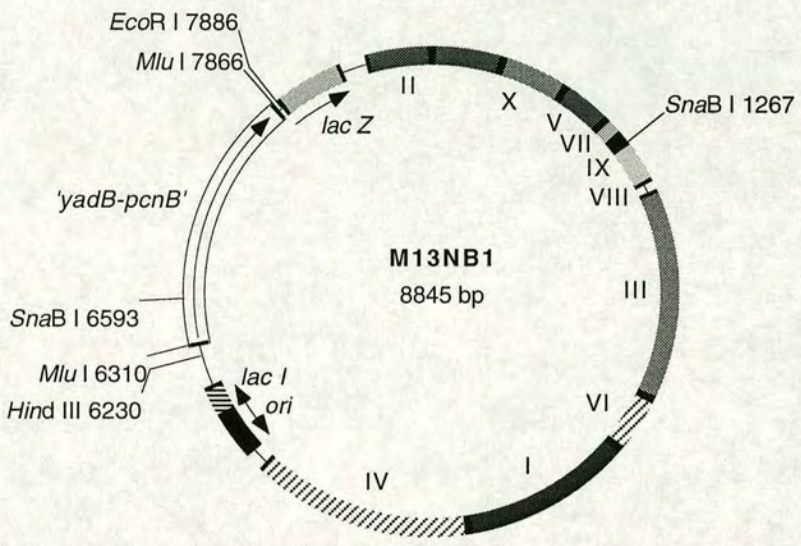
The position of the initiation codon empirically determined in this work is designated as +1. Upstream in-frame codons are indicated by a minus sign. Thus, I shall show that the previously proposed UUG start-codon is located at position -7.

4.3.1 Construction of selected *pcnB*(am) mutants

The 1.6 kb *Mlu* I fragment from pJM513 (Figure 4.2.1) was cloned into the corresponding unique site in M13BM21, a derivative of M13mp19 with an expanded polylinker, creating M13NB1 (Figure 4.3.1). The orientation of the insert was determined by restriction analysis. A number of amber mutations were introduced between the previously proposed start site (Leu_{-7}) and the first known translated codon identified by N-terminal sequencing (Lys_{11}). The location of these mutations (with respect to the start codon subsequently identified by this work) and the previously proposed UUG start codon are

shown in Figure 4.3.2. The associated sequencing autoradiograms in Figures 4.3.3 and 4.3.4 document the introduction of these mutations.

Figure 4.3.1 M13NB1.



M13NB1. Based on a derivative of M13mp19. See text for construction details.

Figure 4.3.2 Location of amber mutations introduced into the sequence surrounding the proposed translational initiation region of *pcnB*.

```
          -7                -2  -1  +1   2       4                   10  11
...UGAUGU UUG ACACUACCGAGG UGU ACU AUU UUU ACC CGA GUCGCUAAUUUUUGC CGC AAG...
          L                 C   T   I   F       R                   R   K
```

Location of introduced amber codons in the *pcnB* translational initiation region. These mutations (coding triplet to TAG) were individually introduced into separate vectors by the Kunkel (1985) method of site-directed mutagenesis. Underlined nucleotide triplets (as mRNA) indicate the sites of the introduced amber codons. Amino acids are identified by their single-letter code below the sequence. Numbers above the sequence indicate the triplet co-ordinates with respect to the translational start codon, designated +1, identified in this work.

Figure 4.3.3 Autoradiograms of dideoxy-sequence ladders confirming the presence of the amber mutations introduced into the sequence upstream from codon +1, in the translational initiation region of *pcnB*.

Autoradiograms of dideoxy-sequence ladders confirming the presence of the amber (TAG) mutations introduced into the sequence upstream from codon +1, in the translational initiation region of *pcnB*. For ease of location and comparison, the ladders have been arranged so that the ATTs (codon +1) from each gel are horizontally aligned. The possible significance of the distinctive run of T residues is discussed in section 5.2.3. Single-stranded template of M13NB1 DNA (bearing the mutated *Mlu* I fragment) was sequenced from oligonucleotide M7537 (5′-d[CGTGCCCGAGTCGGC]-3′). Amber termination codons were introduced by site-directed mutagenesis with the oligonucleotides N7529 (Leu -7) 5′-d[GAATGATGTTAGACACTACCG AGG]-3′, N7915 (Cys -6) 5′-d[CTACCGAGGTAGACTATTTTTACC]-3′ and P8944 (Thr -1) 5′-d[CCGAGGTGTTAGATTTTTACCCG]-3′.

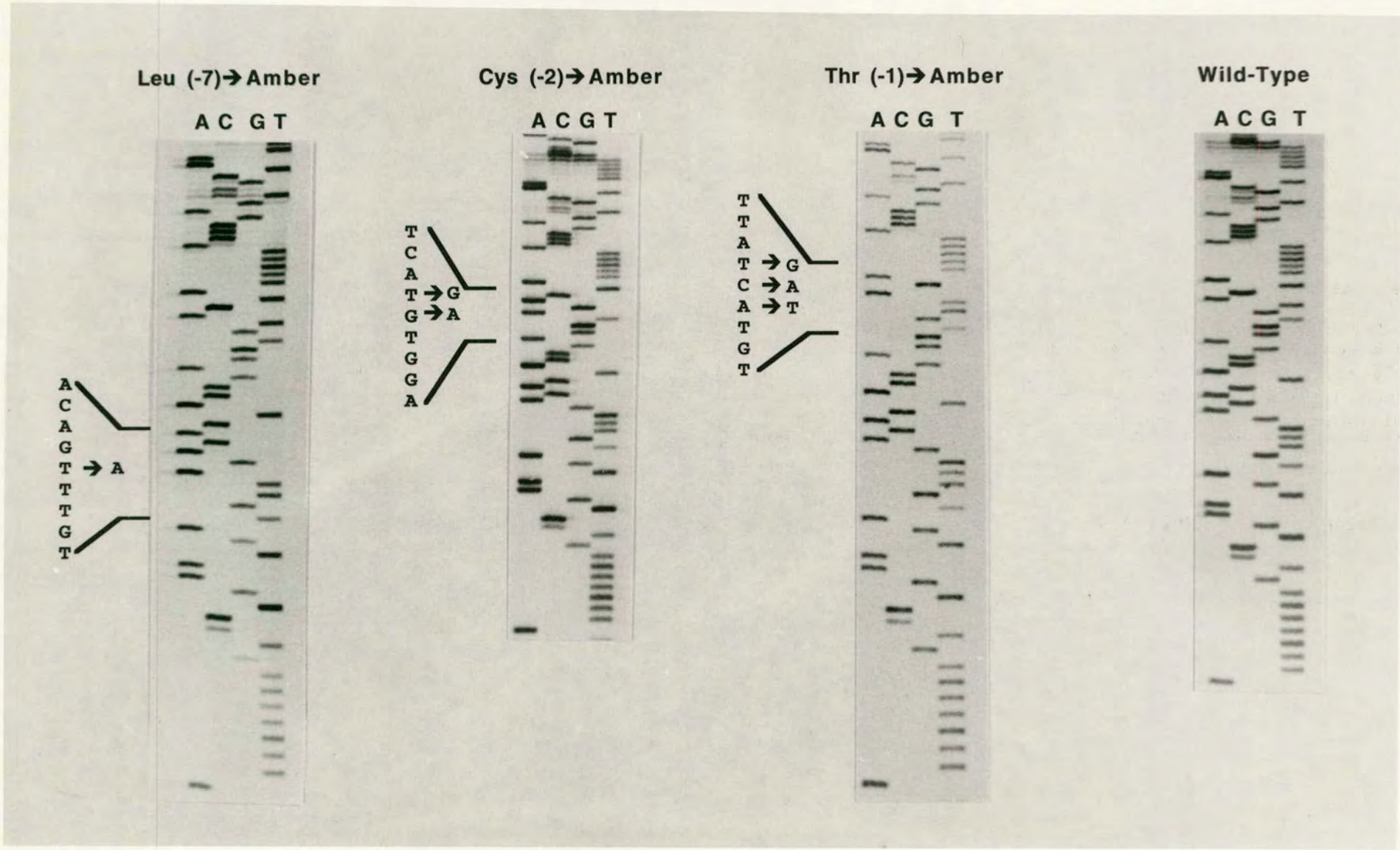
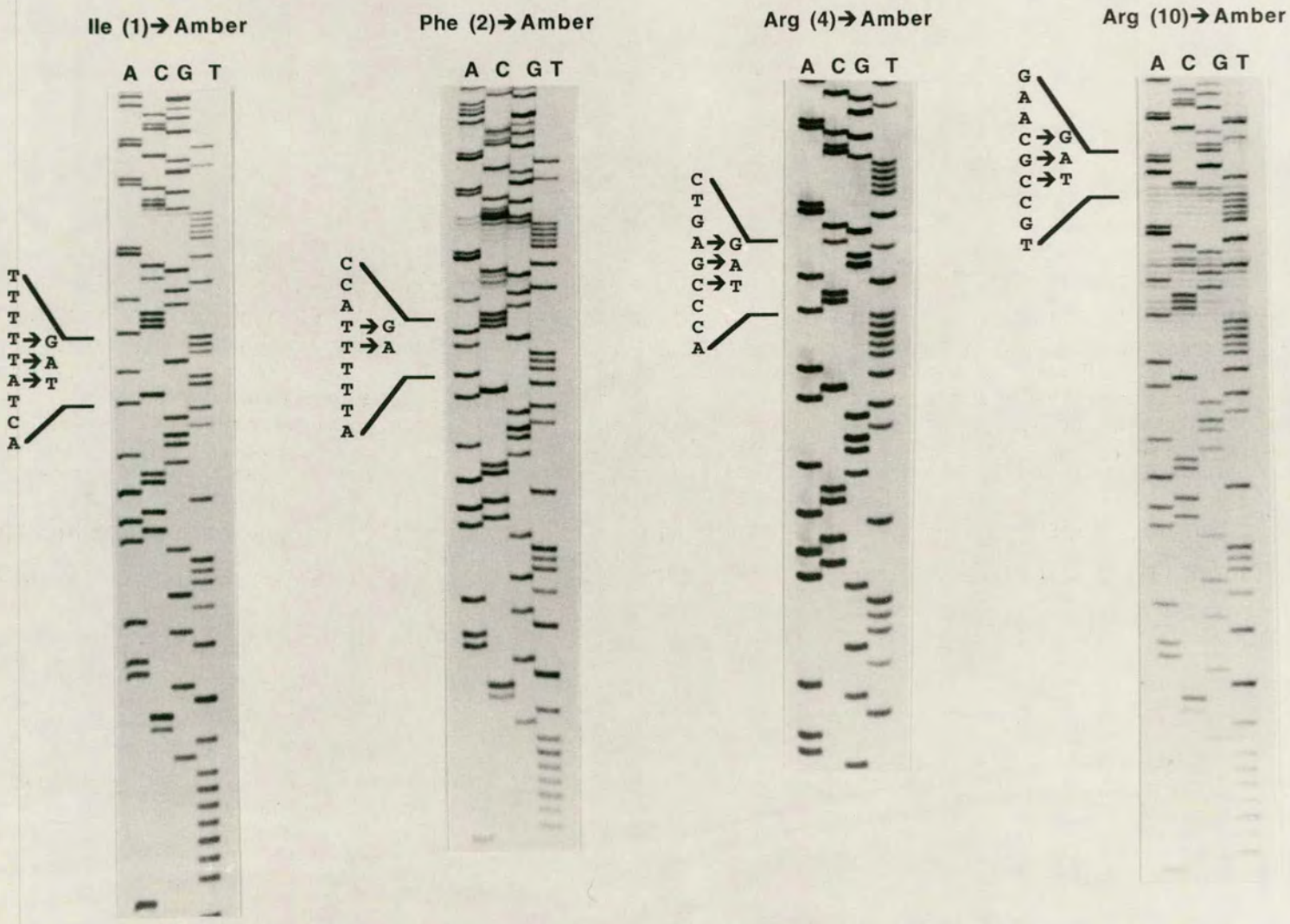

Leu (-7)→Amber
A C G T
A C A G T T G T
→ A
Cys (-2)→Amber
A C G T
T C A T G T G G A
→ G
→ A
Thr (-1)→Amber
A C G T
T T A T C A T G T
→ G
→ A
→ T
Wild-Type
A C G T

Figure 4.3.4 Autoradiograms of dideoxy-sequence ladders confirming the presence of the amber mutations introduced into the sequence downstream from (and including) codon +1 in the translational initiation region of *pcnB*.

Autoradiograms of dideoxy-sequence ladders confirming the presence of the amber mutations introduced into the sequence downstream from (and including) codon +1 in the translational initiation region of *pcnB*. For ease of location and comparison, the ladders have been arranged so that the ATTs (codon +1) from each gel are horizontally aligned. The possible significance of the distinctive run of T residues is discussed in section 5.2.3. Single-stranded template of M13NB1 DNA (bearing the mutated *Mlu* I fragment) was sequenced from oligonucleotide M7537 (5′-d[CGTGCCCGAGTC GGC]-3′). Amber termination codons were introduced by site-directed mutagenesis with the oligonucleotides P8943 (Ile 1) 5′-d[GAGGTGTACTTAGTTTACCCGAG]-3′, S3761 (Phe 2) 5′-d[GAGGTGTACTATTTAGACCCGAG]-3′, S8432 (Arg 4) 5′-d[TTTACCTAGGTCGCTAAT TTTTG]-3′ and N7530 (Arg 10) 5′-d[TAATTTTTGCTAGAAGGTGCTAAG]-3′.


Ile (1)→Amber
A C G T
T T T T A T C A
T→G
T→A
A→T
Phe (2)→Amber
A C G T
C C A T T T T A
→G
→A
Arg (4)→Amber
A C G T
C T G A G C C C A
A→G
G→A
C→T
Arg (10)→Amber
A C G T
G A A C G C C G T
A→G
G→A
C→T

4.3.2 Translation of the *pcnB* message commences from the unusual AUU initiation codon - *in vitro* evidence

In order to determine the effect of the different amber mutations on translation, the mutated *Mlu* I fragments were subcloned into pNB22, a derivative of pJM513 with the corresponding *Mlu* I fragment deleted and the ends re-ligated. The resulting series of plasmids encoding reconstituted *pcnB*(am) were designated pJM513 with a suffix (based on the single-letter amino acid code) indicating the substituted residue. Thus, a UAG replaces the cysteine at codon -2 in pJM513C-2. The presence of the engineered amber codon was confirmed by double-stranded sequencing. These seven constructs were the templates used to direct *in vitro* translation from the encoded *pcnB*(am) sequence. The autoradiograms from this study (Figures 4.3.5 and 4.3.6) appear to show that translation originates from the AUU at position +1, as translation is blocked only by amber codons introduced at this site or downstream from it. The only chromosomal gene in *E. coli* that has previously been reported to initiate translation from an AUU is *infC*, which encodes the translational initiation factor IF3. The significance of this will be discussed later in association with other experimental data.

4.3.3 A Shine-Dalgarno sequence is functionally spaced from the AUU codon

A potential ribosome binding site is located six nucleotides upstream from the AUU in *pcnB* mRNA (Figure 4.3.7). This region of complementarity between the pre-initiation region of the *pcnB* mRNA and the 3′ end of 16S rRNA extends over eight nucleotides, though it is probable that base pairing between these two RNAs is restricted to the first five (i.e. GAGGU). This is because the cytosine in 16S rRNA (indicated by * in figure 4.3.7) that is complementary to the guanine in *pcnB* mRNA may be sequestered by a hairpin structure formed due to the self-complementary nature of the 3′ end of 16S rRNA. However, this putative translational initiation region scores poorly when analysed with weight matrix W101 that represents the sequence around prokaryotic ribosome binding sites (Stormo *et al*, 1982). Weight matrix W101 was derived from a gene set of 124 known starts in an mRNA library by an artificial intelligence algorithm. It is a two dimensional array of values that represents the score for finding each nucleotide, at each position

Figure 4.3.5 *In vitro* translation of PcnB from pJM513 and its *pcnB*(am)-encoding derivatives - 1.

In vitro translation of PcnB from pJM513 and its *pcnB*(am)-encoding derivatives. DNA of each plasmid was prepared and purified on a CsCl gradient. Translated proteins were detected by autoradiography after being labelled with [^{35}S]-methionine and separated on 12% polyacrylamide gels. *pcnB*(am)-encoding derivatives of pJM513 are designated by the single-letter code for the substituted amino acid and a number indicating the residue's position in the open reading frame. SfsA and DksA are proteins encoded by genes located upstream from *pcnB* in the *E. coli* chromosome. The 9 kDa band is believed to be a Cmp fusion protein encoded at the boundary between the insert and the vector (March *et al*, 1989). Bla, β-lactamase and Cmp, chloramphenicol acetyl transferase, are encoded by pBR328, the vector backbone of pJM513. Cmp was partially deleted from pJM513 during its construction.

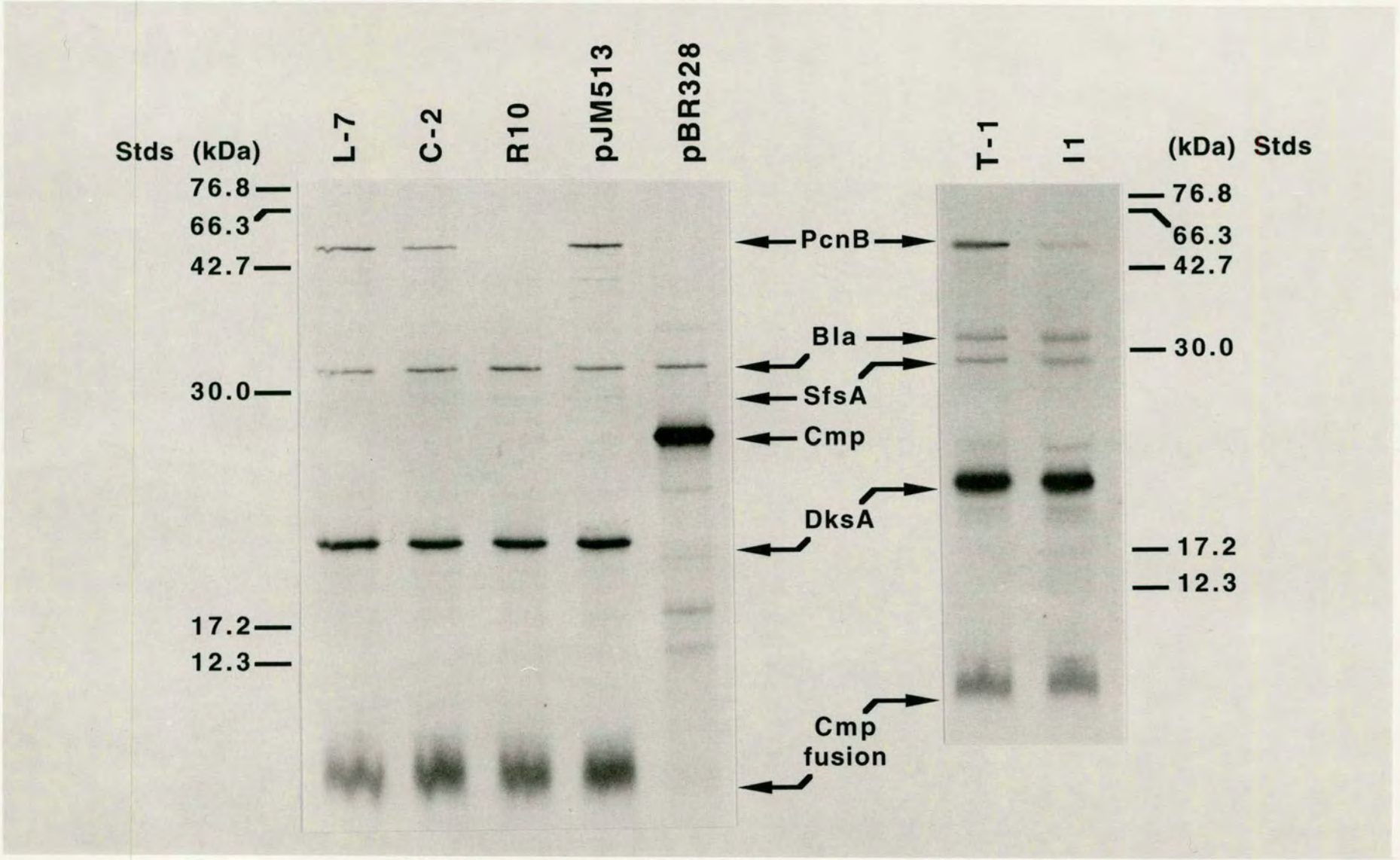
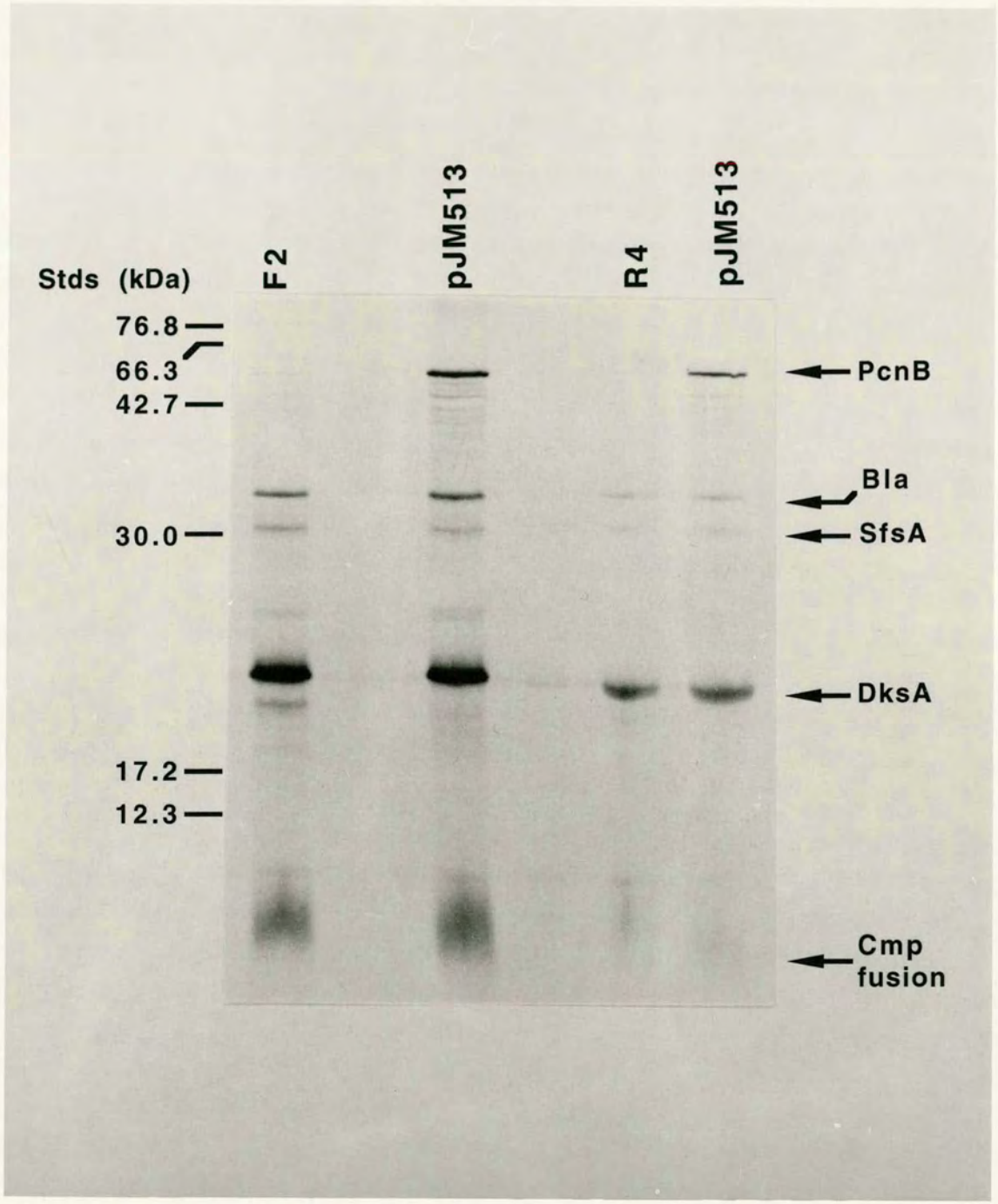

Stds (kDa)
L-7
C-2
R10
pJM513
pBR328
T-1
I1
(kDa) Stds
76.8
66.3
42.7
30.0
17.2
12.3
PcnB
Bla
SfsA
Cmp
DksA
Cmp fusion

Figure 4.3.6 *In vitro* translation of PcnB from pJM513 and its *pcnB*(am)-encoding derivatives - 2.

In vitro translation of PcnB from pJM513 and its *pcnB*(am)-encoding derivatives. DNA of each plasmid was prepared and purified on a CsCl gradient. Translated proteins were detected by autoradiography after being labelled with [^{35}S]-methionine and separated on 12% polyacrylamide gels. *pcnB*(am)-encoding derivatives of pJM513 are designated by the single-letter code for the substituted amino acid and a number indicating the residue's position in the open reading frame. SfsA and DksA are proteins encoded by genes located upstream from *pcnB* in the *E. coli* chromosome. The 9 kDa band is believed to be a Cmp fusion protein encoded at the boundary between the insert and the vector (March *et al*, 1989). Bla, β-lactamase, is encoded by pBR328, the vector backbone of pJM513. Cmp was partially deleted from pJM513 during its construction.


Stds (kDa)
F2
pJM513
R4
pJM513
76.8
66.3
42.7
30.0
17.2
12.3
PcnB
Bla
SfsA
DksA
Cmp
fusion

in the ribosome binding site, within a region of 101 nucleotides. Evaluating whether any particular site is predicted to be a translation initiation site or not is simply done by adding up the corresponding elements of the weight matrix. The fact that the start site of *infC* was not part of the gene set used to train the matrix-generating algorithm (its sequence was unknown at the time) could account for the low score obtained for *pcnB* with this matrix. However, a more likely explanation for the low score is that most computer modeling attempts to identify the typical features of genes. Such approaches often do not perform well on special cases like *infC* and *pcnB*.

Figure 4.3.7 Translational initiation region of *pcnB*.

```
                        SD      Ile                                  Lys
...ugauguuugacacuaccGAGGUGuacuAUUuuuacccgagucgcuaauuuuugccgcaag...
                   ||  |||||*
            3'- AUucCUCCAC -5' 16S rRNA
```

The translational initiation region of *pcnB*. SD, Shine-Dalgarno sequence; Ile, identified initiation codon - this is located at nucleotides 46-48 downstream from the 5' end of the *pcnB* transcript; vertical lines indicate nucleotides with complementarity to the 3'-end of 16S rRNA; * indicates complementary nucleotide pair that may not participate in stabilising the interaction between the mRNA and a 30S ribosomal subunit. See text for details. Lys, first amino acid in protein as determined by N-terminal sequencing of PcnB; underlined nucleotide triplet is the most upstream termination codon in frame with the coding sequence of the gene.

It is interesting to note that although, as expected, pJM513F(2), -R(4) and -R(10) do not support PcnB synthesis, low levels of full-length PcnB are clearly visible in the translation products from pJM513I(1). Work carried out by van der Laken *et al*, 1979 could provide an explanation. In one experiment, they found that UUU can be recognised by initiator tRNA. In *pcnB*, this particular codon immediately follows the AUU initiation codon and lies within functional distance of the Shine-Dalgarno sequence. In addition, toeprint data show that a stable ternary complex can be assembled on a UUU codon (Hartz *et al*, 1988). In their experiment, the toeprint was shifted downstream by one triplet to 18 nucleotides from the initiation codon when $tRNA^{Phe}$ was substituted for the $tRNA_f^{Met}$. An interval of 15 nucleotides typically separates the first base of the initiation codon from the toeprint with most mRNAs.

4.3.4 Translation of the *pcnB* message commences from the unusual AUU initiation codon - *in vivo* evidence

The *in vitro* experiments described in section 4.3.2 helped identify the translational start of *pcnB* by physical means: for each amber mutant (with the exception of I1), a radiolabelled protein of the expected molecular weight was either apparent or absent. These results were supported by data that demonstrated the biological activity of the *pcnB*(am) mutants *in vivo*. This was carried out by measuring the single-cell ampicillin resistance (proportional to gene dosage and hence plasmid copy number) of a Sup^0 Δ*pcnB* strain bearing one of the seven *pcnB*(am) mutant plasmids, compared to an isogenic strain with an appropriate cloned amber suppressor in a separate but compatible vector.

4.3.5 Construction of *pcnB*(am) encoding plasmids with a pMB1-compatible replicon

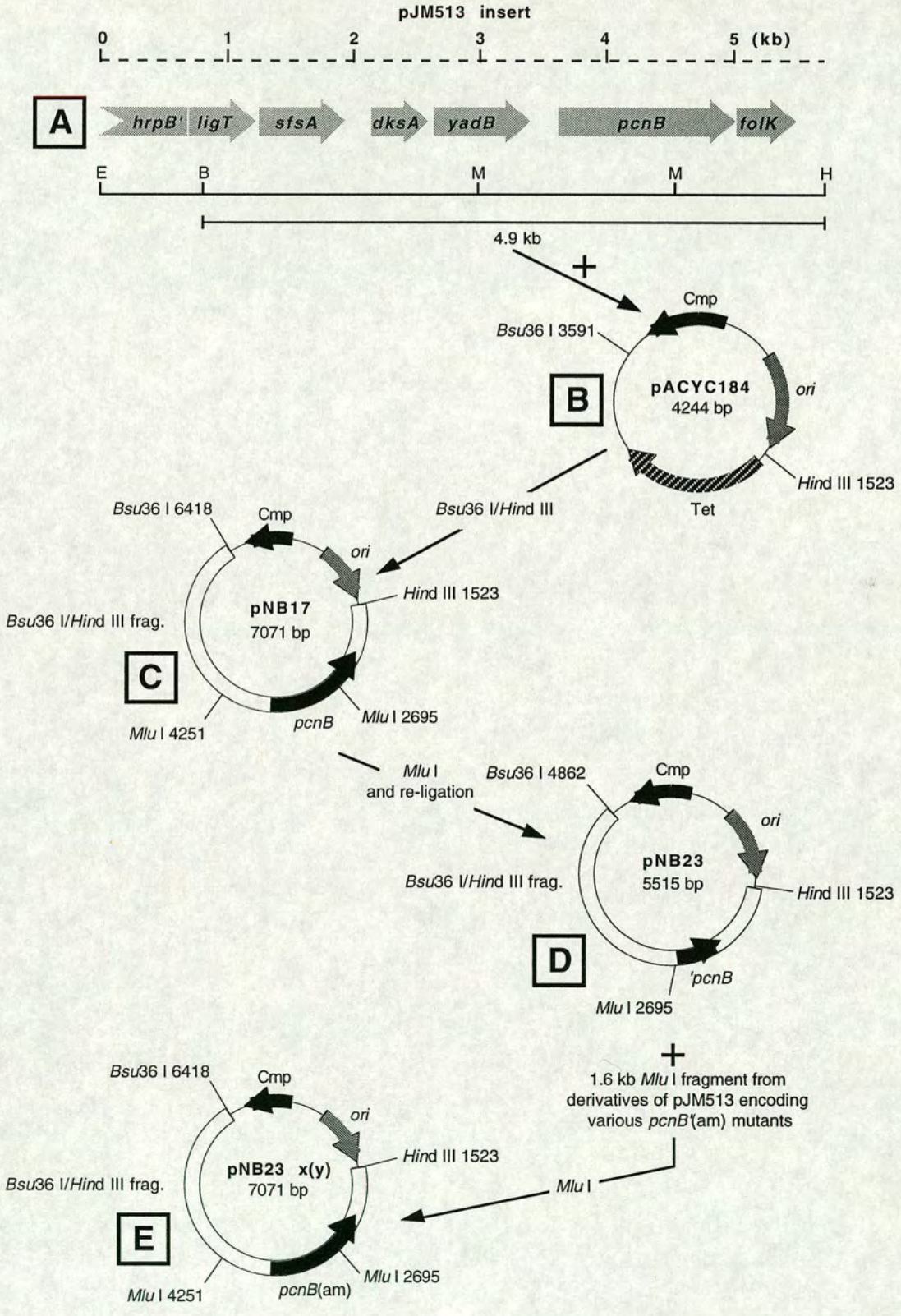
The 1.6 kb *Mlu* I fragments (encoding the 5′ end of *pcnB*) from three of the pJM513 derivatives (encoding various *pcnB*(am) mutations: C-2, -F2 and -R4) were sub-cloned into the *Mlu* I site in pNB23 (encoding the 3′ end of *pcnB*, Figure 4.3.8 D) such that the full coding sequence of *pcnB* was restored in each case. The p15A-derived origin of replication in pNB23 is compatible with the pMB1-derived origin of replication encoded by the amber suppressor plasmids, allowing them to be stably maintained together. Restriction analysis confirmed that the orientation of the *Mlu* I fragments restored the *pcnB* coding sequence in plasmids pNB23C(-2), -F(2) and -R(4) (Figure 4.3.8 E).

4.3.6 An amber mutation in the *pcnB* open reading frame at position C(-2) complements a Sup^0 Δ*pcnB* strain, whereas those at F(2) and R(4) do not

The Sup^0 strain (*sup*$^+$, *ara*, Δ(*lac-pro*), *gyrA*, *argE*(am), *rpoB*, *thi*) supplied with the Interchange™ *in vivo* Amber Suppression Mutagenesis System (Promega) was transduced to Δ*pcnB* Kan^R with a lysate made on strain IR8903, creating Sup^0K. This strain was transformed with pBR322 (suppressor-free and with the same replicon as the suppressor-encoding vectors) and with each *pcnB*(am)-encoding plasmid. The plasmids pCys, pPhe and pArg encoding the eponymous amber suppressors from the Interchange™ system were

Figure 4.3.8 Construction of pMB1-compatible derivatives of pJM513 encoding various *pcnB*(am) mutations.

Construction of pMB1-compatible derivatives of pJM513 encoding various *pcnB*(am) mutations. Diagram A represents the wild-type pJM513 insert. Diagram B is the new vector backbone with a pMB1-compatible p15A origin of replication. Diagrams C and D indicate the intermediary cloning steps. Diagram E represents the final pMB1-compatible *pcnB*(am)-encoding plasmid, where x represents the single-letter code of the substituted amino acid, and y indicates its position in the protein with respect to the initiation codon. See text for additional construction details.


pJM513 insert
0
1
2
3
4
5 (kb)
A
hrpB'
ligT
sfsA
dksA
yadB
pcnB
folK
E
B
M
M
H
4.9 kb
+
Cmp
Bsu36 I 3591
B
pACYC184
4244 bp
ori
Hind III 1523
Tet
Bsu36 I/Hind III
Bsu36 I 6418
Cmp
ori
Hind III 1523
pNB17
7071 bp
Bsu36 I/Hind III frag.
C
Mlu I 4251
pcnB
Mlu I 2695
Mlu I
and re-ligation
Bsu36 I 4862
Cmp
ori
pNB23
5515 bp
Bsu36 I/Hind III frag.
Hind III 1523
D
'pcnB
Mlu I 2695
+
1.6 kb Mlu I fragment from
derivatives of pJM513 encoding
various pcnB'(am) mutants
Mlu I
Bsu36 I 6418
Cmp
ori
Hind III 1523
pNB23 x(y)
7071 bp
Bsu36 I/Hind III frag.
E
Mlu I 4251
pcnB(am)
Mlu I 2695

individually transformed into Sup^{0}K, already harbouring their cognate *pcnB*(am)-encoding plasmid: pCys with pNB23C(-2) and so on.

Evidence for the retention or loss of *in vivo* PcnB activity is presented in Table 4.3.1. Sup^{0}K bearing either pNB23F(2)/pBR322 or pNB23R(4)/pBR322 retains its *pcnB* phenotype, indicating that pNB23F(2) and -R(4) are unable to complement the *pcnB* deletion by supplying biologically active PcnB *in trans*. Wild-type PcnB activity is restored when the cognate amber suppressor is supplied on a second plasmid. These results indicate that the amber codons introduced into these constructs are responsible for the premature termination of *pcnB* translation. The reversion to *pcnB*$^+$ caused by the presence of the suppressor means that the F(2)am and R(4)am mutations are located intragenically. The initiation codon must, therefore, lie upstream from these mutations. In addition to restoring wild-type ampicillin resistance, the amber suppressor plasmids facilitated the production of a cellular excess of PcnB that is responsible for the appearance of the associated diminished colony size phenotype. pNB23C(-2) appears able to supply full-length PcnB *in trans* without requiring its cognate amber suppressor, indicating that the C(-2)am mutation is located 5′ to the initiation codon. Collectively, these results support the conclusions derived from the *in vitro* mapping data described in section 4.3.2.

Table 4.3.1 Ability of various plasmid-borne *pcnB*(am) mutants to complement a defined chromosomal deletion of *pcnB* in a *sup*^{0} background.

Plasmids	Number of colonies at an ampicillin concentration (μg ml^{-1}) of:			
	50	150	500	800
pBR322/C(-2)	112	115	94	102
pCys/C(-2)	95	94	105	100
pBR322/F(2)	104	73	0	0
pPhe/F(2)	110	89	89	95
pBR322/R(4)	106	56	0	0
pArg/R(4)	98	91	109	92
pBR322/pNB23	103	68	0	0
pBR322/pNB17	120	95	108	94

Fresh transformants of Sup^{0}K were grown exponentially in LB-Cmp (25 μg ml^{-1}) and LB-Amp (50 μg ml^{-1}) to an OD_{540} of 0.3, diluted in LB broth so as to obtain approximately 200 colonies per plate on antibiotic-free medium and spread on LB plates containing Cmp (25 μg ml^{-1}) and Amp at the concentrations indicated. The numbers of colonies on plates without ampicillin were defined as 100 (i.e. they are a '100% growth' control); the figures in each horizontal row are percentages of these initial values.

4.3.7 The translational block imposed by a UAG codon at position I(1) in the *pcnB* open reading frame is unaffected by the presence of an amber suppressor

The I(1)am mutation was also subjected to essentially the same *in vivo* suppressor-reversion study. In this case, a chromosomally encoded suppressor (*supP*) was used that inserts a leucine residue at the site of an amber codon. This was the closest and most conservative substitution available: leucine and isoleucine are hydrophobic amino acids with similar sized aliphatic side chains. The SupP strain (Promega Interchange™ system) was transduced to Δ*pcnB* Kan^R with a lysate made on strain IR8903, creating SupPK. This strain and Sup^{0}K were both transformed with pJM513I(1).

SupPK was also transformed separately with pJM513 and pNB22 as *pcnB*$^+$ and *pcnB*$^-$ controls, respectively.

Table 4.3.2 Ability of a plasmid-borne *pcnB*(am) mutant that encodes an amber termination triplet in place of its wild-type initiation codon to complement a defined chromosomal deletion of *pcnB* in a *sup*^{0} or *supP* background.

	Number of colonies at an ampicillin concentration (μg ml^{-1}) of:				
Strain	50	150	500	800	Plasmid
Sup^{0}K	103	79	0	0	pJM513I(1)
SupPK	95	77	0	0	pJM513I(1)
SupPK	109	100	108	102	pJM513
SupPK	99	70	0	0	pNB22

Fresh transformants of Sup^{0}K and SupPK were grown exponentially in LB-Amp (50 μg ml^{-1}) to an OD_{540} of 0.3, diluted in LB broth so as to obtain approximately 200 colonies per plate on antibiotic-free medium and spread on LB plates containing ampicillin at the concentrations indicated. The numbers of colonies on plates without ampicillin were defined as 100 (i.e. they are a '100% growth' control); the figures in each horizontal row are percentages of these initial values.

The amber codon inserted at position I(1) appears to prevent the *pcnB*(am)-encoding plasmid pJM513I(1) from complementing the Δ*pcnB* deletion. However, unlike the situation with pNB23F(2) and -R4, the presence of an amber suppressor charged with a leucine residue (supplied by strain SupPK) does not appear to exert any effect on ampicillin resistance and hence does not facilitate the expression of biologically active PcnB (Table 4.3.2). Colonies of SupPK harbouring pJM513I(1) are of normal size, which is also indicative that *pcnB* is not expressed from this construct. This behaviour is consistent with the belief that I(1) is the site of the initiation codon. Specific differences in the primary (and therefore probably secondary) structure of the RNA moiety uniquely distinguish fMet-tRNA$_f$ from elongator tRNAs. Some of these differences ensure that tRNA$_f^{Met}$ cannot be used in elongation - the

absence of base-pairing between the first base from the 5′ end and another base within the acceptor arm, the formylation of methionine which blocks its free -NH_2 group (Wu and RajBhandary, 1997) - others, such as the presence of a series of three G-C base pairs in the stem of the anticodon arm, ensure that fMet-$tRNA_f$ is the only aminoacyl-tRNA that can directly enter the partial-P site of the initiation complex (Seong and RajBhandary, 1987). Thus it was anticipated that an amber suppressing elongator tRNA would be incapable of reverting the translational block imposed by a UAG substitution at the site of the initiation codon. The same result could be obtained, of course, if the introduced amber codon disrupted either the Shine-Dalgarno sequence or other uncharacterised, but important, upstream contextual information. However this is unlikely here because the engineered C(-2)am codon, located only two triplets upstream from the identified I(1) initiation codon, does not prevent the synthesis of full-length PcnB.

4.4 IF3-mediated control of *pcnB* expression

The role of AUU as a critical determinant in the autogenous control of cellular IF3 levels and its rarity as an initiation codon is well documented (Gold *et al*, 1984). Because the translation of *pcnB* also appears to be initiated from the same triplet, it was considered that IF3 might also influence its expression. To investigate this possiblity, the β-galactosidase activity from the *pcnB'-'lacZ* translational fusion, λNB25 (section 4.2.1), in an *infC* background, was compared with its activity in an isogenic *infC*$^+$ strain.

4.4.1 The β-galactosidase activity of a PcnB'-'LacZ translational fusion is derepressed in an *infC* background

The *infC* allele used for this investigation, *infC362*, was obtained from JK378 (Sussman *et al*, 1996). Strains bearing this mutation grow slightly slower on rich and minimal media and exhibit somewhat increased levels of intact IF3 (from Western blot analysis). They do not show temperature sensitivity. The presence of this allele was confirmed by measuring the level of *infC*-specific derepression of an *infC'-'lacZ* fusion in the *infC362* background. JK378 and its isogenic *infC*$^+$ strain, JK382, were infected with λRS544, also from R. W. Simons. This vector carrys the mini-*infC'-'lacZ* fusion with the native *infC* AUU start codon altered to CUG. Sussman *et al*, 1996 assign CUG to the same class of IF3-sensitive initiation codons as AUU. Transcription is under the control of P_{tac}. Lysogens of λRS544 were prepared and selected as described in section 2.4.2, because this construct lacks the selectable kanamycin resistance gene present in λNB21 and λNB25. A similar level of *infC*-specific derepression to that reported in Sussman *et al*, 1996 was obtained with JK378/λRS544 lysogens (approximately 3-fold).

To facilitate meaningful comparisons with other studies in this lab and earlier data from this work, MM38 was again chosen as the genetic background for studies to establish the β-galactosidase activity from λNB25. The *infC362* allele from JK378 was transferred by P1 transduction to MM38/λNB25 to create MM378/λNB25. The co-transduction frequency between *infC* and *zdh*::Tn*10* in JK378 is about 60%. For the *infC*$^+$ control, MM38/λNB25 was transduced with a lysate made on the isogenic *infC*$^+$ strain, JK382 to create MM382/λNB25. The *infC362* and *infC*$^+$ alleles were similarly introduced into MM38/λNB21. This was done in order to

demonstrate that like *infC*, the expression of *pcnB* is unaffected by IF3 at the level of transcription. The behaviour of the *infC362* allele in MM38 was also compared with published data for the mutation in the donor strain by transducing MM38/λRS544 with P1 lysates made on JK378 and JK382. The β-galactosidase activity detected in these lysogens is shown in Table 4.4.1.

Table 4.4.1 β-galactosidase activity of λNB25, a *pcnB'-'lacZ* translational fusion vector in isogenic *infC* and *infC*$^+$ strains.

	Background (strain)		
Lambda-based *lacZ* fusion	*infC*$^+$ (MM382)	*infC362* (MM378)	Derepression
pcnB Translational (λNB25)	77 ± 1.3	298 ± 6.6	3.9
pcnB Transcriptional (λNB21)	53 ± 0.7	58 ± 0.9	1.1
infC Translational (λRS544)	20 ± 0.6	63 ± 0.9	3.2

β-galactosidase assays were carried out according to Miller (1972) in the indicated strain, harbouring *lacZ* fusions based on a monolysogenic phage lambda (λRS45). Values (in Miller Units) for each construct were obtained from cells growing exponentially in LB broth + 0.2% glucose (and, where appropriate, Kan at 50 μg ml^{-1}) at 37°C and are the average of three measurements ± the standard deviation.

The acquisition of the *infC362* allele by MM38 leads to an approximately four-fold increase in β-galactosidase activity from the *pcnB'-'lacZ* translational fusion, whereas no significant derepression is observed with the *pcnB-lacZ* transcriptional fusion. These results suggest that the *in vivo* expression of *pcnB* is regulated, at least in part, by the activity of IF3 at the level of translation. The level of *infC*-dependent derepression obtained for λRS544 in the MM378 background was comparable to the published data for the mini-*infC'-'lacZ* fusion in JK378.

4.4.2 A cellular excess of IF3 restores the β-galactosidase activity of a PcnB'-'LacZ translational fusion in an *infC* background to the wild-type level

To demonstrate that the derepression observed in MM378/λNB25 (*infC362*) is *infC*-specific, β-galactosidase levels were measured in this strain whilst harbouring an IF3 overproducing plasmid. If IF3 levels negatively control the translation of *pcnB* mRNA, then a cellular excess of IF3 should cause repression of *pcnB* expression at the translational level *in vivo*. The plasmid, pSB1 (Butler *et al*, 1986) is a derivative of pBR322 in which only the *infC* gene is inserted. It overexpresses IF3 eight-fold from the native proximal *infC* promoter, pO'.

Table 4.4.2 The effect of IF3 overexpression on the β-galactosidase activity of λNB25, a *pcnB'-'lacZ* translational fusion vector, in isogenic *infC* and *infC*$^+$ strains.

Background (strain)	pBR322 (vector plasmid)	pSB1 (*infC* plasmid)	Repression
infC$^+$ (MM382/λNB25)	77 ± 1.1	72 ± 1.3	0.07
infC362 (MM378/λNB25)	298 ± 5.7	79 ± 1.4	3.8

β-galactosidase assays were carried out according to Miller (1972) in the indicated strain, harbouring *lacZ* fusions based on a monolysogenic phage lambda (λRS45). Values (in Miller Units) for each construct were obtained from cells growing exponentially in LB broth + 0.2% glucose (and, where appropriate, kanamycin at 50 μg ml^{-1}) at 37°C and are the average of three measurements ± the standard deviation.

In the *infC362* background, the IF3 overexpressing plasmid pSB1 represses the expression of the *pcnB'-'lacZ* translational fusion carried by λNB25 approximately four-fold, restoring its β-galactosidase activity to the level seen when in a wild-type background (Table 4.4.2). Together with the data in Table 4.4.1, this indicates that translation of *pcnB* mRNA specifically responds to changing levels of IF3 *in vivo*. The expression of the PcnB'-'LacZ hybrid from λNB25 in wild-type cells is not significantly repressed by a cellular excess of IF3. These results appear to show that a single chromosomal copy of wild-type *infC* is able to induce maximal IF3-mediated

(dependent) repression of *pcnB*. In contrast, a 2.5-fold repression of β-galactosidase activity from an *infC'-'lacZ* translational fusion in an *infC* wild-type strain was observed when a cellular excess of IF3 was provided by pSB1 (Butler *et al*, 1986). The *pcnB* transcript does not appear to share any of the regions of 16S rRNA complementarity that *infC* mRNA possesses (except the AUU and Shine-Dalgarno sequence) and which are believed to stabilise the initiation complex, facilitating IF3-independent translation of IF3. The absence of similar regions of stabilising complementarity in *pcnB* mRNA could account for the maximal repression of *pcnB* occurring at lower IF3 levels than IF3 itself responds to.

4.4.3 Changing the *pcnB* AUU initiator codon to AUG in a *pcnB'-'lacZ* translational fusion in order to abolish IF3 control of *pcnB* expression led to the insert encoding the fusion becoming unstable

As *infC* and *pcnB* share the same unusual initiation codon and since this codon has been show to play an important role in *infC* autoregulation (Butler *et al*, 1987), it is probable that the AUU codon in *pcnB* is also a critical *cis*-acting element through which the negative regulatory control exerted by IF3 on *pcnB* expression at the level of translation is mediated. It has previously been shown that *infC* autoregulation can be abolished by changing its AUU initiation codon to AUG (Butler *et al* 1987). This method was also employed in an attempt to uncouple *pcnB* expression from IF3 control. The I(1) codon of *pcnB* carried by M13NB1 was mutagenised to AUG forming M13NB1I(1)M (Figure 4.4.1).

The mutated 1.6 kb *Mlu* I insert was gel-purified to free it from 'contaminating' M13 DNA and cut with *Eco*R V. The 775 bp *Eco*R V/*Mlu* I cleavage product was end-filled using the Klenow fragment of DNA Polymerase I and cloned in to the *Bam*H I site (also end-filled with Klenow) of pRS552 creating what should have been pNB26. Apart from the ATT to ATG change and 172 bp absent from the 5' end of the cloned insert, the resultant (ATG)*pcnB'-'lacZ* translational fusion is identical to that encoded by pNB25. A new wild-type (ATT)*pcnB'-'lacZ* fusion (called pNB27) was also constructed in order that it contain the same truncated insert as pNB26. When blue transformants (on medium containing X-gal) of these clones were checked, it was found that only the wild-type 775 bp insert had been cloned (pNB27); no (ATG)*pcnB'-'lacZ* fusions had been generated. Although it was

Figure 4.4.1 Autoradiogram of dideoxy-sequence ladder confirming the presence of the ATT to ATG mutation introduced into the sequence at codon +1 in the translational initiation region of *pcnB*.

Autoradiogram of dideoxy-sequence ladder confirming the presence of the ATT to ATG mutation introduced into the sequence at codon +1 in the translational initiation region of *pcnB*. The possible significance of the distinctive run of T residues is discussed in section 5.2.3. Single-stranded template of M13NB1 DNA (bearing the mutated *Mlu* I fragment) was sequenced from oligonucleotide M7537 (5′-d[CGTGCCCGAGTCGGC]-3′). Initiation site conversion had been introduced by site-directed mutagenesis with the oligonucleotide S3762 (Ile 1 Met) 5′-d[GAGGTGTACTATGTTTACCCGAG]-3′.

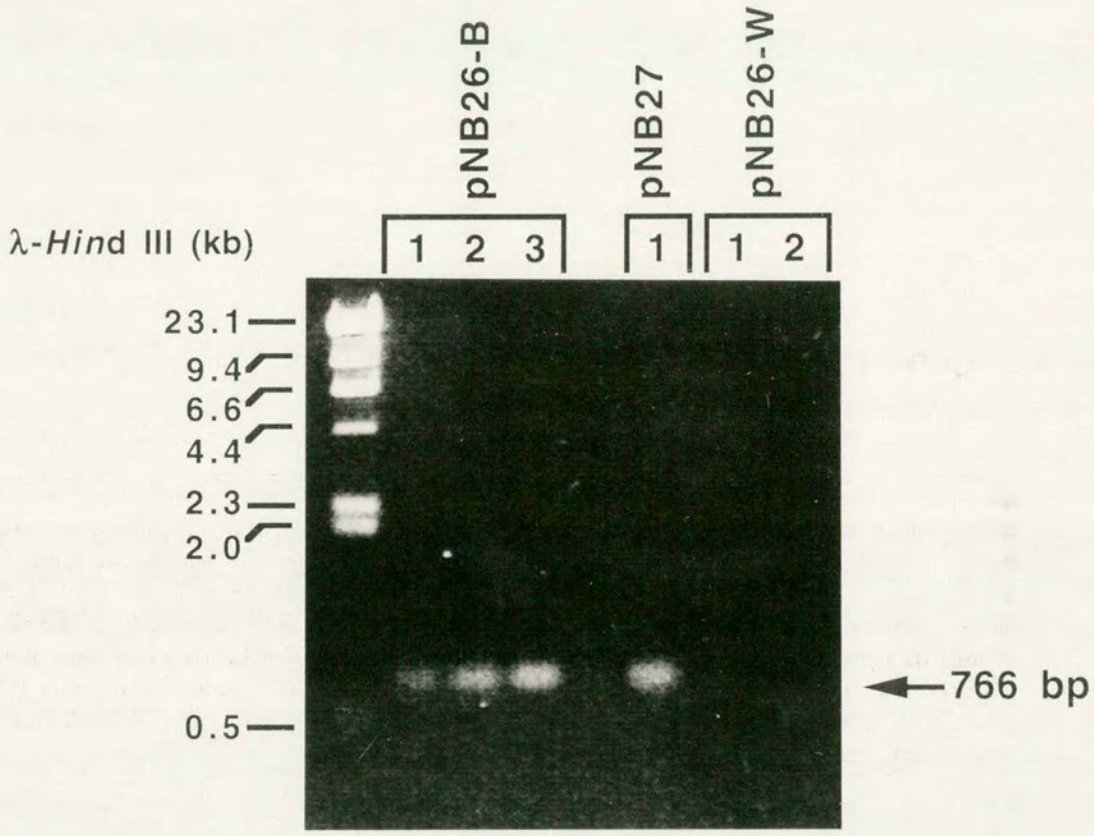

Ile (1)→Met
A C G T
T T T T T A T C A
T → G

anticipated that full-length *pcnB* with an ATG start codon on a multicopy vector might be difficult to clone, the expected fusion product would only contain the N-terminal 43 amino acids from PcnB. A repeat cloning produced the same result. However, after overnight incubation at 4°C, a blue putative pNB26 clone spread to obtain discrete colonies gave rise to predominantly white colonies. The constructs present in these cells were named according to the colour of their transformants: pNB26-B if blue, pNB26-W if white. Cells bearing pNB27 (wild-type fusion) remained entirely blue when similarly treated. The minority blue progeny (pNB26-B-bearing cells) of this blue-to-white conversion also gave rise to a mixed population of white and blue colonies , indicating that the characteristic is heritable. These observations suggest that the single base change from ATT to ATG causes an instability leading to the complete or partial loss of the insert (and some of the backbone vector DNA, as an *Eco*R I site in it is no longer detectable). The ratio of white to blue clones increased during overnight incubation in broth prior to plasmid DNA mini-preparation: cells diluted from such a culture produced about 200 white colonies for every blue colony. These blue transformants were shown to harbour the desired insert by amplifying it directly from the colony with PCR, thus circumventing the need to purify the plasmid by overnight growth in broth. Three independently isolated blue pNB26-B transformants each produced a PCR product of the expected size (of 766 bp), indicating the presence of the cloned insert in the desired orientation. No amplified product was obtained from white progeny of blue pNB26-B clones (Figure 4.4.2).

This procedure gave the same result when repeated using one of the PCR-confirmed blue pNB26-B clones as the source colony. Ultimately, the (AUG) *pcnB'-'lacZ* translational fusion was to be transferred to phage λ where its β-galactosidase activity as a monolysogen of an *infC*$^-$ strain could be compared with its activity in an *infC*$^+$ strain. Once transferred to phage λ, it was hoped that the unstable fusion in the multicopy vector would aquire stability as a monolysogen. Instead of transforming NM621/λRS45 with purified pNB26 (Section 3.5.2), a single blue colony derived directly from a colony shown to harbour pNB26-B was boiled (for five minutes followed by a short centrifugation) and used as the source of the plasmid DNA. A number of pale-blue Kan^r colonies were obtained on X-gal/Kan plates. Several of these were subjected to colony-PCR. Unfortunately, this yielded no amplification

Figure 4.4.2 Confirmation of cloning the (ATG)*pcnB'-'lacZ* translational fusion construct by colony-PCR.

Confirmation of cloning the (ATG)*pcnB'-'lacZ* translational fusion construct by colony-PCR. Three independently isolated blue colonies (on medium containing X-gal) transformed with pNB26, a plasmid intended to bear an (ATG)*pcnB'-'lacZ* translational fusion, were subjected to colony-PCR with the oligonucleotides 763K (anneals to *pcnB*) 5'-d[TCATGGACCAGACA A]-3' and M7062 (anneals to *lacZ*) 5'-d[GGTAACGCCAGGGTTTTCCC]-3'. The size of the amplified product is indicated by the arrow (766 bp). pNB26-B forms blue colonies in cells grown on medium containing X-gal; pNB26-W forms white colonies in cells grown on medium containing X-gal; pNB27 is the wild-type equivalent of pNB26-B. See text for construction details.


pNB26-B
pNB27
pNB26-W
λ-Hind III (kb)
1 2 3
1
1 2
23.1
9.4
6.6
4.4
2.3
2.0
0.5
766 bp

products of the expected size (data not shown). The procedure was repeated with the same outcome. It is assumed that the insert was lost during homologous recombination with the lysogenised phage λ; recombination must have occurred as NM621/λRS45 had become kanamycin resistant (whilst remaining ampicillin sensitive). It is possible that the *recA*$^+$ status of NM621 (essential for the recombination process) may have made the insert encoding the (AUG)*pcnB'-'lacZ* fusion even more unstable; the original blue pNB26-B clones were obtained by transforming DH5α, a *recA*$^-$ strain, with the ligation mixture. This recombination deficiency might explain why the clone was originally isolated. If it is assumed that a DNA instability is created by changing the *pcnB* initiation codon from AUU to AUG, the question posed by the experiment might be answered by mutating the AUU to either GUG or UUG. Both these latter codons have been assigned to the same class as AUG on the basis of their translation efficiency and absence of derepression in an *infC* background (Sussman *et al*, 1996). If the instability of the insert was due to the production of a lethal excess of the PcnB'-'LacZ hybrid protein, replacing the AUG with these other start codons could significantly reduce expression from the fusion; β-galactosidase activity from GUG and UUG mini-*infC'-'lacZ* fusions is approximately 7-8-fold lower than with AUG in an *infC*$^+$ strain (Sussman *et al*, 1996). An alternative approach to reducing expression (if this is the cause of the instability) would be to replace the native *pcnB* promoter with a tightly controllable one such as P_{trp} or P_{ara}.

4.5 Summary

The expression of *pcnB* does not appear to be autogenously regulated at the level of translation. Using a *pcnB'-'lacZ* translational fusion, this was demonstrated by the same procedure employed to investigate whether *pcnB* is subject to autoregulation at the level of transcription (Section 3.5). The *pcnB* translational initiation codon was shown to be AUU. Until this work, *infC*, which encodes IF3, was the only chromosomally encoded gene known to commence translation from this codon. Evidence that *pcnB* utilises this unusual initiation codon was obtained by introducing amber termination codons at various positions (one per *pcnB*-encoding plasmid) within and outside the putative open reading frame, and testing their ability to block the full-length synthesis of PcnB by *in vitro* coupled transcription/translation. This evidence was supported by complementation studies on Δ*pcnB* strains harbouring the various *pcnB* (am)-encoding plasmids. A Shine-Dalgarno sequence is located within functional distance from the AUU codon. Like *infC*, *pcnB* expression was found to be negatively regulated by IF3 levels. Expression of a *pcnB'-'lacZ* translational fusion is derepressed four-fold in an *infC* background. A cellular excess of IF3, provided on an overexpressing plasmid, specifically restores the β-galactosidase activity of this fusion to the wild-type level. These results appear to show that a single chromosomal copy of wild-type *infC* is able to induce maximal IF3-mediated (dependent) repression of *pcnB*. An attempt to uncouple *pcnB* expression from IF3 negative regulation, by altering the AUU initiation codon to AUG in a *pcnB'-'lacZ* translational fusion, led to the insert encoding the fusion becoming unstable and unclonable.

IF-3, growth rate and plasmid copy number

A concomitant elevation in the co-ordinate expression of initiation factors and ribosomes is observed with an increase in the growth rate of *E. coli*. Modulation of the IF-3/ribosome ratio associated with this growth rate-dependent regulatory system provides the control signal for the autogenous expression of *infC* through its AUU initiation codon. (See section 1.4.3 for a more detailed description). As *pcnB* has been shown to utilise an AUU triplet as an initiation codon, its expression is subject to regulation by IF-3 (Section 4.3.2) and presumably, therefore, to changes in cellular growth rate. PcnB is a

positive regulator of ColE1-type plasmid copy number. It has been known for some time that the copy number of this class of plasmids decreases with increasing growth rate (Lin-Chao and Bremer, 1986). It is therefore possible that the reduction in plasmid copy number associated with an elevation in growth rate is related to IF3-dependent repression of *pcnB* expression.

However, because of the steady-state ratio of IF-3/ribosomes (1:7), it is predicted that the expression of *pcnB* will shadow the expression of IF-3. That is, the *pcnB*/ribosome ratio will also be essentially constant (though the value of this ratio is unknown). As the number of ribosomes increases with growth rate, the cellular level of IF-3 will rise so that the IF-3/ribosome ratio is maintained. The expression of *pcnB* will be coordinately regulated with IF-3 levels, because the absolute number of IF3-free 30S subunits, which is a function of the IF-3/ribosome ratio, will increase. As a consequence, *pcnB* expression would be expected to rise, accompanied, according to this scenario, by an elevation in plasmid copy number. The degree to which the latter occurs would depend upon whether a threshold level of *pcnB* exists, above which there would be no increase in plasmid copy number; the existence of *pcnB* point mutants (such as *pcnB21*) that support copy numbers below wild-type but above those observed in *pcnB* deletion mutants indicates that there is probably a lower threshold level of *pcnB*, below which plasmid copy number starts to fall.

In Section 4.4.2, it was shown that at fast growth rate (exponential in rich medium - LB broth), *pcnB* expression is maximally repressed by a single chromosomally-encoded copy of wild-type *infC*, as the presence of an IF-3 overproducing plasmid (pSB1) has little effect on the expression of *pcnB*. Under these conditions, however, it is possible that pSB1 (which is a pBR322-derivative and therefore itself subject to a reduction in copy number with increasing growth rate and sensitive to *pcnB* levels) is at low copy number and therefore the *infC* gene-dosage is lower than expected. As a result, the level of IF-3 expression from pSB1 may be much lower than the eight-fold overproduction claimed, without supporting data, by Butler *et al*, 1986. (The creators of pSB1 do not show the data pertaining to the overexpression of IF-3 from this plasmid, nor do they indicate how and under what conditions the IF-3 level was assayed). Under the experimental conditions employed in Section 4.4.2, therefore, pSB1 might be unable to repress *pcnB* expression below the level observed in an *infC*$^+$ strain. A potential excess of IF-3 from

pSB1 might also contribute to a reduction in its own copy number. Fewer IF3-free 30S subunits available for initiating the expression of *pcnB* might depress the cellular level of PcnB. In time, an equilibrium between plasmid copy number (and hence *infC* gene-dosage), IF-3 levels and *pcnB* expression would be attained. This situation could be resolved by cloning *infC* into a vector, such as pWSK29 (Wang and Kushner, 1991), whose origin of replication (derived from pSC101) is insensitive to *pcnB*. Perhaps, at slow growth rates, when the copy number of pSB1 is higher, expression of *pcnB* would be repressed (by IF-3) below the level observed in an *infC*$^+$ strain with a single chromosomal copy of *infC*. This may indeed be the case, as pSB1 represses the expression of an *infC'-'lacZ* translational fusion in an *infC*$^+$ strain by approximately 2.5-fold below the wild-type level (Butler *et al*, 1986) when grown in a nutrionally poor medium (MOPS) with glucose as the sole carbon source (again, it should be noted that the authors do not indicate at what temperature this assay was carried out). In addition to the DNA sequence fused to the reporter gene, many other factors could influence the results of this assay, including strain differences and temperature.

The possibility that a decreased level of *pcnB* expression (caused by an increased cellular level of IF-3) is the sole cause of reduced plasmid copy number at higher growth rates would need to be confirmed experimentally. An *infC* mutant (defective in non-canonical initiation codon discrimination) harbouring a ColE1-type plasmid (e.g. pBR325) would be assayed for plasmid copy number at different growth rates. This would be done either by sampling at various points throughout the growth curve from lag to stationary phase in a single type of medium, or by sampling plasmid copy number in exponential phase during steady-state growth in media of varying nutritional richness (e.g. minimal-glycerol to LB broth). Plasmid copy number could be determined by a DNA-DNA hybridisation method (Klotsky and Schwartz, 1987). If variable *pcnB* expression (coupled to growth rate through IF-3 levels) determines plasmid copy number, then under these experimental conditions, plasmid copy number should be constant at all growth rates.

CHAPTER 5

Analysis of the sequence and genetic organisation of *pcnB* homologues in eubacteria other than *E. coli*

5.1 Identification of *pcnB* homologues in other eubacteria that potentially could utilise an AUU triplet as a translation initiation codon.

5.1.1 Introduction

Mutational analysis (discussed in Chapter 4) of *E. coli pcnB* has revealed that its translation is initiated from an AUU codon (Figure 5.1.1). The only other gene in the *E. coli* chromosome known to use AUU as a translational start codon is *infC*, which encodes IF3 (Initiation Factor 3). By re-examining the appropriate region of nucleotide sequence, it was possible to determine whether *pcnB* homologues (and the closely related *cca* gene, tRNA nucleotidyl transferase) from eubacterial species other than *E. coli* might be able to utilise an AUU triplet as an initiation codon. The utilisation of AUU (by a gene) as an initiation codon is significant, as it has been shown to bestow negative regulatory control by IF3 on that gene (Sussman *et al*, 1996 and this work). IF3 also negatively autoregulates *infC*. In addition, a more fundamental question can be asked: is an AUU triplet utilised as an initiation codon more often than presently believed? Since the initiation codon of a gene is rarely identified by empirical means, the frequency of AUU (or other rare codon) utilisation may be underestimated. However, if *pcnB* and *infC* are the only genes in *E. coli* to employ an AUU initiation codon, then the specific control of *pcnB* translation by IF3 has to be explained in terms of the biological function of PcnB and is therefore interesting.

Figure 5.1.1 Translational initiation region of *E. coli pcnB*

```
Opa    Leu          SD        Ile                                  Lys
UGAuguUUGacacuaccGAGGUguacuAUUuuuacccgagucgcuaauuuuugccgcAAGgugcua
```

SD, predicted Shine-Dalgarno sequence; Ile, initiation codon; Opa, termination codon; Leu, originally proposed initiation codon; Lys, amino-terminal residue identified by N-terminal sequencing.

5.1.2 Approach

The nucleotide sequence surrounding the proposed translational initiation region of those *pcnB* homologues from other eubacteria exhibiting the greatest similarity to *E. coli pcnB* were examined for the presence of an in-frame ATT codon separated by an appropriate distance from a Shine-Dalgarno sequence. The full extent of the open reading frame, in the 5′ direction, was determined for each *pcnB* homologue. Starting from the upstream stop codon, the amino acid sequence was examined in a 3′ direction for the presence of isoleucine, methionine, valine and leucine residues. Their location was used as a point-of-reference for identifying the presence of in-frame ATT, ATG, GTG and TTG codons, respectively. The upstream nucleotide sequence preceding each occurrence of these candidate initiation codons was examined by eye for a potential Shine-Dalgarno sequence (optimally) located between 5 and 13 nucleotides from the query initiation codon.

5.1.3 Analysis of *pcnB* homologues with a potential AUU initiation codon

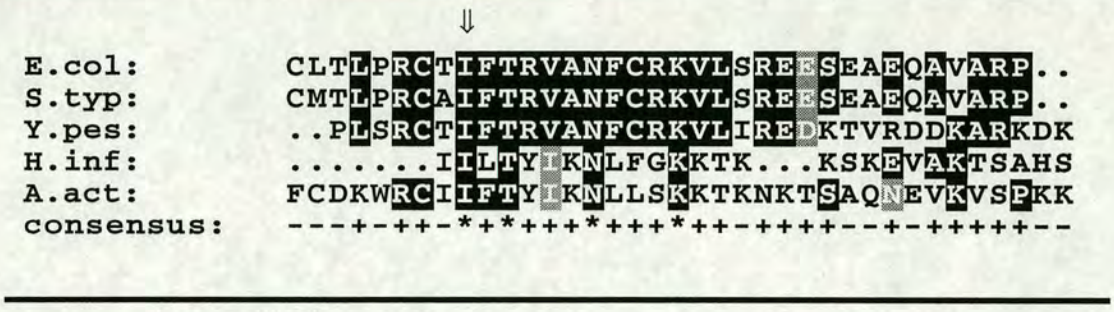
Four eubacterial *pcnB* homologues in addition to *E. coli pcnB* were found to possess potential translational initiation regions that could utilise an AUU start codon (Figures 5.1.2 and 5.1.3). The nucleotide alignment spans the region from the most upstream termination codon in frame with the coding sequence of each gene, to the triplet coding for the first amino acid (lysine) in *pcnB* from *E. coli*, as determined by N-terminal sequencing. The included sequence from the *pcnB* homologue in *Haemophilus influenzae*, was necessarily extended in a 5′ direction from the termination codon (located just three nucleotides upstream from the proposed AUU initiation codon) in order to reveal a potential Shine-Dalgarno sequence. The proposed initiation codon of the *pcnB* homologue from *Actinobacillus actinomycetemcomitans* is AUC. Although there are no reports of this triplet being used as an initiation codon in a native context, evidence exists to suggest that translational initiation can occur from this codon (Sussman *et al*, 1996; Cheng *et al*, 1994; Romero and Garcia, 1991). When the start site of *tnp* (IS*10* transposase) was changed from its native AUG to AUC, it was found that this triplet was able to support translation, albeit at a reduced level. In addition, the level of expression from

Figure 5.1.2 Alignment of predicted translational initiation regions from eubacterial homologues of *E. coli pcnB* that potentially could utilise an AUU triplet as a translational initiation codon.

```
                                SD        IC                                    Lys
E.col:  ugauguuugacacuaccGAGGUguacuAUUuuuacccgagucgcuaauuuuugccgcAAGgug

S.typ:  uaauguaugacacuaccGAGGUgcgcuAUUuuuacccgagucgcuaauuuuugccguAAGgug

Y.pes:        uagccacuaucGAGGUguaccAUUuuuacccgaguagccaauuucugccguAAGgug

H.inf:           aacaaaacGGAGuaaauuAUUcuuacuuauaucaaaaauuuguuuggcAAGaag

A.act:  uauuuuugugauaaauGGAGauguauuAUCuuuaccuauauuaaaaaucuuuugagcAAGaaa
                              SD        IC                                    Lys
```

E.col, *Escherichia coli*; S.typ, *Salmonella typhi*; Y.pes, *Yersinia pestis*; H.inf, *Haemophilus influenzae*; A.act, *Actinobacillus actinomycetemcomitans*; SD, Putative Shine-Dalgarno sequence; IC, putative initiation codon; Lys, first amino acid in protein as determined by N-terminal sequencing of the *pcnB* from *E. coli*; underlined nucleotides are the most upstream stop codons in frame with the coding sequence of each gene. The potential coding sequence of *pcnB* in *A. actinomycetemcomitans* extends upstream for a further 15 codons.

Figure 5.1.3 Alignment of predicted N-terminal regions from eubacterial homologues of *E. coli pcnB* that potentially could utilise an AUU triplet as a translational initiation codon.



Alignment of predicted N-terminal regions from eubacterial homologues of *E. coli pcnB* that potentially could utilise an AUU triplet as a translational initiation codon. The alignment was made with Clustal X. Arrow indicates predicted position of formyl-methionine incorporation. Identical residues are boxed in black with a *, conserved residues are boxed in black with a +, similar residues are boxed in grey with a + and residues at a position below the conservation threshold are plain with a - (threshold of 60%, at each position). E.col, *Escherichia coli*; S.typ, *Salmonella typhi*; Y.pes, *Yersinia pestis*; H.inf, *Haemophilus influenzae*; A.act, *Actinobacillus actinomycetemcomitans*.

the AUC-modified gene increased when assayed in an IF-3 mutant, mirroring the behaviour of *infC* and *pcnB* from *E. coli* as well as *tnp* with a non-native AUU initiation codon (Sussman *et al*, 1996). Although there is an in-frame AUU triplet immediately upstream from the AUC codon, the sub-optimal spacing between it and the proposed Shine-Dalgarno sequence makes it a less attractive candidate (this factor applies equally to *pcnB* in *H. influenzae*, where a second AUU is located in tandem with the proposed initiation codon). In addition, the second codon appears to be an almost conserved UUU triplet (phenylalanine); translational initiation of *A. actinomycetemcomitans pcnB* from the AUU at position -1 would disrupt this conservation. It should be noted that the *pcnB* open reading frame in *A. actinomycetemcomitans* extends for 24 codons upstream from the proposed AUC start site.

5.2 Identification of the promoters from *pcnB* homologues with a potential AUU translational initiation codon

5.2.1 Introduction

A search was also carried out to locate potential promoters for those *pcnB* homologues - identified in section 5.1.3 - that may utilise an AUU triplet as a translational initiation codon. Although the number of nucleotides separating the first transcribed base from the ribosome binding site can vary widely from gene to gene (in the range of a few bases to tracts of more than a thousand nucleotides), it was hoped that identification of the promoters would complement the evidence presented for the initiation codon data.

5.2.2 Approach

Primer extension analysis has revealed that the distance between the -10 region of the σ^{70} promoter and the ATT initiation codon in the *pcnB* transcript from *E. coli* is 52 nucleotides (section 3.2.5). Based on this information, the region of DNA up to 200 nucleotides 5′ of the proposed initiation codon was examined for potential σ^{70} promoters. The first sequence examined was from *H. influenzae*. A near consensus -10 and -35 region is located 106 nucleotides upstream from the ATT initiation codon. Examination of the sequence flanking the candidate promoter revealed a run of six T residues separated from the 3′-end of the -10 region by 9 nucleotides. A distinctive run of seven Ts, with identical spacing from the -10 region, is present in *pcnB* from *E. coli*. Assuming this sequence determinant to be significant, the promoters from the remaining three species were rapidly located by scanning the sequence for a -10 region proximal to a run of thymine bases.

5.2.3 Analysis of the promoters from *pcnB* homologues

An alignment of these predicted promoters (Figure 5.2.1) reveals that although the number of Ts in the run varies from five to seven, the nine nucleotide spacing is conserved in all the species examined. When immediately preceded by a stem-loop, a short tract of poly(T) is a characteristic of rho-independent transcriptional terminators. However, none

Figure 5.2.1 Alignment of predicted σ^{70}-like promoters from eubacterial homologues of *E. coli pcnB* that potentially could utilise an AUU triplet as a translational initiation codon.

```
                    -35                    -10
E.col: tgtaaattcaacaTTCTCAaatgcgtcatgctgagcTATGATtagccgctaTTTTTTTgtcctgaatgatgtttgacactaccgaggt
S.typ: tgtaaatccggcaTTCTCAaatgcgtctcgctgagcTATGATtagccgctaTTTTTgctcatcgtctttcgcattgctgcggcgttaa
Y.pes: aaacacaacggcaTTCTCAaagtaagtatggtgagcTATGATtagccgctgTTTTTctgtcgttctgttatttattagccactatcga
H.inf: gttaattcattctaTTGAAAaatacaaaagtgcggtTAAAATtaaccgcacTTTTTTgcatttaaaaatgcaggaaaatccacttgcc
A.act: tgggttaatcctgCTAATTcattgataaaagtgcggTCAAAAaacacattaTTTTTcgaccgcactttcactttatttttgtgataaa
                    -35                    -10
```

E.col, *Escherichia coli*; S.typ, *Salmonella typhi*; Y.pes, *Yersinia pestis*; H.inf, *Haemophilus influenzae*; A.act, *Actinobacillus actinomycetemcomitans*; -35 and -10 indicate the eponymous regions of consensus in *E. coli* σ^{70} promoters.

of the poly(T) tracts aligned in Figure 5.2.1 appear to be preceded by an appropriately spaced stem-loop.

An alternative explanation for the run of thymine bases is that the expression of *pcnB* is subject to a phenomenon known variously as reiterative transcription, transcriptional slippage, pseudo-templated transcription or RNA polymerase stuttering (Jacques and Kolakofsky, 1991). This reaction is characterised by the repetitive addition of nucleotides to the 3'-end of a nascent transcript due to slippage between the transcript and the DNA template. Transcripts produced by this stuttering mechanism are not extended productively to include downstream sequences.

The intracellular concentration of UTP appears to determine whether RNA polymerase enters a reiterative cycle (Liu *et al*, 1994). It has been suggested that the only requirement for UTP-dependent reiterative transcription to occur is a run of at least three T residues (in the non-template strand sequence) located at or very near the beginning of the initially transcribed region of the promoter (Qi and Turnbough, 1995). When UTP levels are high, the weak base-pairing between the run of U residues in the transcript and A residues in the DNA template allows slippage between the two strands, which shifts the transcript one base upstream. RNA polymerase then adds a U residue to the 3'-end of the transcript and the entire process can be repeated, in some cases for more than 50 cycles. The poly(U) transcripts produced by this stuttering reaction fail to undergo normal elongation. The result is a low level of productive transcription and gene expression. However when UTP levels are low, the synthesis of poly(U) tracts in the initially transcribed region is reduced, allowing efficient transcriptional initiation and production of full-length transcript.

Because of the above process, a continuous polymerization reaction occurs, which generates RNA transcripts with long stretches of homopolynucleotide sequence (up to several hundred nucleotides). Polyacrylamide gel electrophoresis can therefore be used to identify whether a promoter is subject to reiterative transcription, since during the reaction a fraction of the synthesised chains is released, giving rise to a regularly spaced ladder of products. The exact mechanism of reiterative transcription is unclear, but it is known that the intensity of the slippage reaction depends on the cellular

concentration of the reiteratively incorporated nucleotide. High concentrations of the reiteratively incorporated NTP cause a large amount of slippage products, while low concentrations of the same NTP diminish the slippage reaction. The efficiency of transcription thus correlates with the cellular NTP concentration.

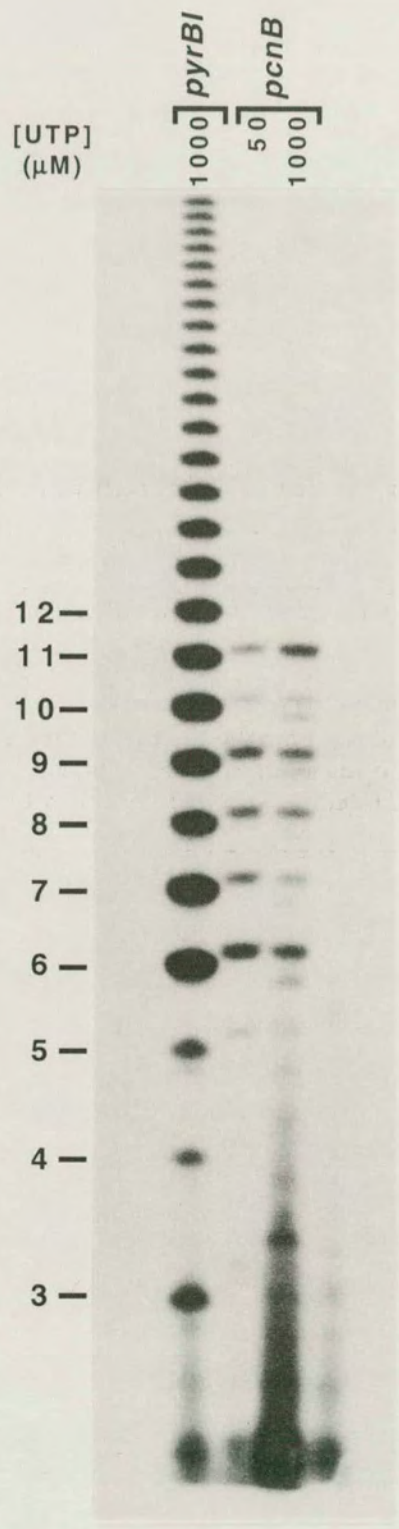
All the examples of regulation by reiterative transcription in *E. coli* involve genes with an obvious connection to pyrimidine metabolism. Examples of these include *pyrBI* encoding aspartate transcarbamylase subunits (Liu *et al*, 1994), *codBA* encoding cytosine permease and cytosine deaminase (Qi and Turnbough, 1995) and *upp* encoding uracil phosphoribosyltransferase (Tu and Turnbough, 1997). Given that *pcnB* plays a role in RNA turnover, a model involving both *pcnB* and *folK* (*folK* is believed to be co-transcribed and translationally coupled to *pcnB*) can be envisaged linking RNA degradation and UTP-mediated regulation. FolK (2-amino-4-hydroxy-6-hydroxylmethyldihydropteridine pyrophosphokinase) catalyses the first step in dihydrofolate biosynthesis. N5, N10-methylenetetrahydrofolate (derived from dihydrofolate) is required for the methylation of dUMP to dTMP. Dividing cells require a supply of dTMP for the synthesis of DNA. If DNA synthesis is inhibited or its rate reduced, this might lead to a build up in UTP. So, elevated levels of UTP might induce UTP-dependent reiterative transcription in the *pcnB* promoter and repress expression of *folK* (and *pcnB*), which is no longer required because the cell (in its current state) has a reduced requirement for dihydrofolate. In addition, the repression of *pcnB* expression itself might be of benefit to the cell in this quiescent state, as it would result in increased stability of mRNA in readiness for rapid expression of necessary proteins when growth conditions were more favourable.

5.2.4 The *E. coli pcnB* promoter, although possessing a run of seven T residues in its initially transcribed region, is not sensitive to UTP-dependent reiterative transcription

To determine whether UTP-dependent reiterative transcription occurs within the *pcnB* initially transcribed region, a DNA template containing the *pcnB* promoter was transcribed *in vitro* in reaction mixtures containing 200 μM each ATP, GTP, and CTP and either 50 or 1000 μM [α-^{32}P]UTP (Qi *et al*, 1996). This work was carried out by Dr. Charles Turnbough at the University of Alabama at Birmingham. Transcripts produced in the *in vitro* reactions were separated on a 25% polyacrylamide sequencing gel and visualised by autoradiography (Figure 5.2.2). At 1000 μM UTP, there is no evidence that reiterative transcription occurs at the *pcnB* promoter. Under the same conditions the *pyrBI* promoter generates a ladder of reiterative transcripts, with the longest transcript containing well over 25 nucleotides. At both concentrations of UTP, the *pcnB* promoter exhibits simple abortive initiation. Abortive initiation is a naturally occurring property of most prokaryotic promoters. It occurs because the initial polymerisation reaction of transcription does not immediately yield long productive transcripts. Instead, short RNAs from two to 10 nucleotides are synthesised and released in several repetitive rounds of reactions (Carpousis and Gralla, 1980). The sequence of short transcripts in an abortive transcript ladder, and hence the transcriptional start site, can be determined by measuring the gaps between the transcripts. Differences in these gaps are due to base-specific effects on electrophoretic mobility. Bases retard transcript mobility in the following order: G>A≥U>C. There is little obvious difference between A and U. However, G addition clearly retards transcript mobility much more than A or U, and C addition clearly retards mobility less than A or U. Inspection of the gaps in lane two (*pcnB* at 50 μM UTP) for transcripts between five and 11 nucleotides in length reveals similar gaps, except for a uniform reduction due to transcript length. This pattern suggests that from base six to base 11, the nucleotides in the transcripts are A or U. If correct, this sequence suggests that these transcripts initiate at position G6 (when numbered from the first base after the -10 hexamer) and the 11-mer ends at the end of the seven T/U tract. This putative G6 start site is the major site of transcriptional initiation from the *pcnB* promoter at both high and low UTP levels. A shift in UTP concentration to 1000 μM promotes the appearance of doublet bands for each

Figure 5.2.2 Effects of severe UTP limitation and excess on the synthesis of transcripts initiated at the *pcnB* promoter.

Analysis of transcripts initiated at the *pcnB* promoter. A DNA template containing the *pcnB* promoter region from *E. coli* was transcribed *in vitro* in reaction mixtures containing the following nucleoside triphosphates: 200 μM each ATP, GTP, and CTP and either 50 μM (lane 2) or 1000 μM [α-^{32}P]UTP (lanes 1 and 3). An autoradiograph is shown of a 25% polyacrylamide gel used to separate the the transcripts. The numbers down the left-hand side indicate the length (in nucleotides) of each transcript. Transcripts initiated from the *pyrBI* promoter are included as a positive control for UTP-dependent reiterative transcription.


[UTP]
(μM)
1000
pyrBI
50
1000
pcnB
12
11
10
9
8
7
6
5
4
3

length of abortive transcript, indicating that under these conditions, a low level of transcription is being initiated from a second, less efficient, start site. Transcripts from this minor start site are probably initiated from position C7, because their faster migration rate indicates a lower G content. Like the transcripts initiated from position G6, these putative C7-initiated transcripts also terminate at the end of the seven T/U tract. Initiation of transcript synthesis from position C7 is likely to be stimulated by elevated UTP levels, because high concentrations of the second nucleotide in a nascent transcript are known to promote initiation at that position (McClure *et al*, 1978; Nierman and Chamberlin, 1979). UTP-sensitive selection of alternative transcriptional start sites probably occurs because of a requirement for rapid formation of a covalent bond between the first and second nucleotides of a transcript to achieve efficient transcriptional initiation. Rapid formation of the first internucleotide bond requires ample concentrations of both nucleotides. A low concentration of the second nucleotide can dramatically reduce initiation efficiency, or it can cause the selection of an alternative start site (at the same promoter) that does not use the limiting nucleotide in the formation of the first internucleotide bond (Sørensen *et al*, 1993). This effect has been seen at many promoters, including *codBA* and *upp*. However in terms of *pcnB* expression, the minor differences in start-site selection were found to have little effect on the level of operon expression (data not shown). This was demonstrated *in vivo* by measuring the β-galactosidase activity from a *pcnB::lacZ* fusion under conditions of pyrimidine excess and starvation. The fusion, as a recombinant phage lambda, was introduced into the chromosome of a pyrimidine auxotroph (*car-94*, which is defective for carbamoylphosphate synthase, the first enzyme of the pyrimidine biosynthesis pathway). The lysogen was grown in glucose-minimal salts medium containing either uracil or UMP as the pyrimidine source, which provides a condition of pyrimidine excess or limitation, respectively. The level of β-galactosidase was assayed as an indicator of *pcnB* expression.

Addendum

Since the submission of this thesis, an alternative explanation has been found that might account for the poly(U) tract located in the initially transcribed region of *pcnB*. A uridine-rich sequence, upstream from the Shine-Dalgarno region of the *rnd* transcript (encoding the tRNA-processing enzyme RNase

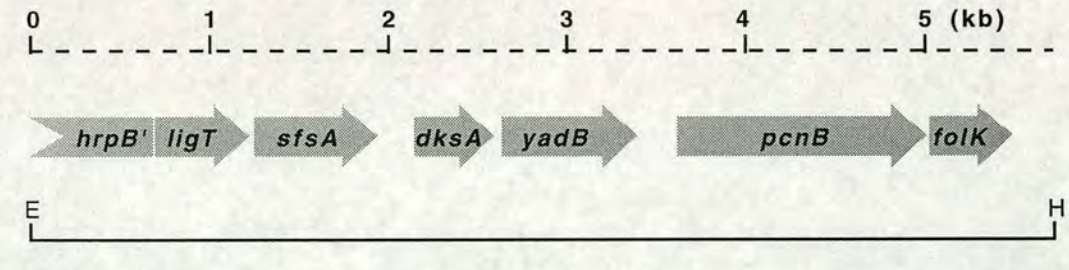
D), has been shown to be essential for the translation of this mRNA (Zhang and Deutscher, 1992). Further evidence presented by these workers suggests that the poly(U) tract, consisting of a run of eight U residues, plays a role in the initial binding of ribosomes to the translation initiation region of the message. In addition to the RNA-RNA interactions between mRNA and ribosomes, it is believed that ribosomal proteins also participate in the selection of translation start sites. Ribosomal protein S1 is known to play a role in translational initiation, being directly involved in recognition and binding of mRNAs by 30S ribosomal particles. Based on UV-crosslinking of protein S1 to phage and bacterial messages *in vitro*, it has been proposed that pyrimidine-rich regions upstream of the Shine-Dalgarno sequence might interact with protein S1 and serve as ribosome recognition sites (Boni *et al*, 1991). Zhang and Deutscher (1992) suggest that the U-rich region may only become necessary for mRNAs that utilise weaker initiation codons (translation of *rnd* commences from a UUG) or weaker Shine-Dalgarno sequences. In these cases, stable binding between ribosomes and the translation initiation region of mRNA may require additional sites for RNA-RNA or RNA-protein interaction. To specifically determine whether the poly(U) sequence in *pcnB* participates in its translational enhancement, residues within this stretch could be changed by site-directed mutagenesis and the effect of these alterations on *pcnB* expression examined.

5.3 Database search for eubacterial species that share with *E. coli* the equivalent genetic organisation flanking their homologue of *pcnB*

5.3.1 Introduction

The *pcnB* gene maps to approximately 3.4 minutes on the *E. coli* chromosome. It is immediately flanked (60 nucleotides) upstream by an uncharacterised open reading frame (924 bases long) designated *yadB* and beyond that a gene called *dksA*. On the 3' side of *pcnB*, its TGA stop codon overlaps with the ATG start codon of *folK* (cgcATGAcag). By following an appropriate search strategy, it was possible to discover whether this genetic organisation in *E. coli* (Figure 5.3.1) is found in other related and unrelated species. Confirmation of similar organisation was made on the basis of sequence similarity or known biological activity. The presence of a *yadB* homologue located upstream from *pcnB* in other species might suggest that *yadB* is biologically functional and that its activity could, in some way, be related to the activity of *pcnB*. In addition, because it is believed that *pcnB* and *folK* are co-transcribed and translationally coupled (Talarico *et al*, 1992), it is possible that their functions are related in some unknown way (PcnB is known to possess a poly(A) polymerase activity and modulates ColE1 plasmid copy number by destabilising the negative regulator RNA I; FolK plays a role in folic acid biosynthesis). The genetic linkage of *folK* to *pcnB* in other species might suggest that the activities of their respective gene products are related enough to require them to be co-expressed.

Figure 5.3.1 Genetic organisation at the 3.4 minute region on the *E. coli* chromosome flanking *pcnB*.



Genetic organisation at the 3.4 minute region on the *E. coli* chromosome flanking *pcnB*. E, *Eco*R I; *Hin*d III Map is drawn to scale.

5.3.2 Approach

In order to ascertain whether homologues of the genes and open reading frame flanking *pcnB* in *E. coli* (Figure 5.3.1) are present and in the same spatial arrangment in other eubacterial species, a survey was made of the eubacterial sequences deposited in the 'Complete and Unfinished Microbial Genomes' database at The National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html). This database contains complete and partial genome sequences originating from whole genome sequencing projects, and is therefore suitable for investigating the genetic organisation of a specific region within the genome. To ensure that the data available from other eubacteria (that are currently not the focus of a whole genome sequencing project and therefore not represented in the Complete and Unfinished Microbial Genomes database) were not missed, the query sequences were submitted in parallel to the eubacterial subset of the GenBank nr database at http://www.ncbi.nlm.nih.gov/blast/. The searches were performed using version 2 of the TBLASTN program which compares an amino acid query sequence against a nucleotide sequence database dynamically translated in all reading frames. The program identifies regions of localised similarity by utilising a gapped local alignment search algorithm (Altschul *et al*, 1997). The gene products translated from *E. coli dksA*, *yadB*, *pcnB* and *folK* were submitted to the two databases. The amino acid sequences of *E. coli cca* (Cudny *et al*, 1986) and *gltX* (Breton *et al*, 1986) were also submitted to the same databases. This was necessary in order to distinguish between eubacterial *cca* enzymes and poly(A) polymerases, which show similarity over a 25 kDa amino-terminal domain (Yue *et al*, 1996). The gene product of *yadB* shows similarity over 25 kDa to the N-terminal half of glutamyl-tRNA synthetase, encoded by *gltX* (Masters *et al*, 1992). A subject sequence producing a significant alignment against *E. coli* PcnB and CCA was classified accordingly by comparing the Expected (E) Values returned with these two query sequences; the subject sequence being grouped with the query sequence returning the lowest E value (i.e. the lowest probability of the alignment occurring by chance alone). The same method was employed to distinguish between the gene products of *yadB* and *gltX*. The spatial arrangment of *dksA*, *yadB*, *pcnB* and *folK* homologues within an organism's genome was determined by examining the nucleotide co-ordinates that

spanned each region of alignment. For an organism whose genome sequence was unfinished, this frequently involved obtaining the relevant unassembled sequences (as contigs) directly from the web site of the institution responsible for the sequencing project and ordering and assembling them.

5.3.3 Analysis of the genetic organisation flanking *pcnB* in other eubacteria

A distillation of the work in Section 5.3.2 is presented in the form of two tables (Tables 5.3.1 and 5.3.2), with the latter showing the E values returned for each search. A taxonomic classification of those eubacteria found to possess a homologue of *E. coli pcnB* is included for reference purposes (Table 5.3.3).

From this database study, it would appear that :-

- Proteobacteria belonging to the β and γ subdivisions encode both *pcnB* and *cca* homologues of *E. coli* genes. In those eubacteria that are more distantly related to *E. coli* - many of which seem to possess only one of these activities - sequence alignment alone cannot reliably distinguish between the two proteins. The recent characterisation of the *papS* gene product in *Bacillus subtilis* illustrates this observation (Raynal *et al*, 1998). PapS returned a significantly better alignment with *E. coli* PcnB than CCA (2e-33 compared with 2e-14), which led to the prediction that this protein was a poly(A) polymerase. However, when cloned and overexpressed, PapS showed no detectable polyadenylation activity, but was found to possess a CCA-adding activity similar to *E. coli* tRNA nucleotidyltransferase. The classification of a gene as either *pcnB* or *cca* from a bacterium listed below *B. subtilis* in Table 5.3.1 should not, therefore, be considered reliable. A line drawn across the table indicates this quasi-threshold. It is probable that a homologue of *cca* is encoded by all eubacteria, irrespective of whether a *pcnB* homologue is present.

- A *folK* homologue is present in all eubacteria found to possess a homologue of *E. coli pcnB*, with the exception of the Spirochaetes and Rickettsia. Also, in all members of the γ subdivision and at least one representative from the β subdivision (*Bordetella pertussis*), *folK* occupies the same relative genetic location (with respect to *pcnB*) identified in *E. coli*. The conservation of this genetic linkage between *pcnB* and *folK*

would seem to suggest that the activities of their respective gene products are related enough to require them to be co-expressed.

- A homologue of *gltX* is encoded by all eubacteria, regardless of whether a *yadB* homologue is present.

- Barring a few unrelated exceptions, homologues of *dksA* are encoded by all eubacteria found to possess a homologue of *E. coli pcnB*. In *Chlamydia trachomatis* and Proteobacteria belonging to the γ subdivision, *dksA* occupies the same relative genetic location (with respect to *pcnB*) identified in *E. coli*.

- A *yadB* homologue is present in representatives from almost all the subdivisions of the Proteobacteria (with the striking exception of *H. influenzae* and *A. actinomycetemcomitans*) that possess a homologue of *E. coli pcnB*. Although *H. influenzae* is closely related to *E. coli* by rRNA sequence comparisons, it has an extremely streamlined genome less than half the size of chromosomes of other proteobacteria (Fleischmann *et al*, 1995). *H. influenzae* may have discarded *yadB* or transferred its function to another polypeptide (*gltX*?) due to the selective pressure that caused the drastic decrease in size of its genome. This would imply that *yadB* may not perform a function absolutely necessary for bacterial survival, but may be conditionally required under (certain) growth conditions encountered by a wide variety of other species. From the data available at the time of the search (February, 1999), it also appears that in all bacteria belonging to the γ subdivision, *yadB* occupies the same relative genetic location (with respect to *pcnB*) identified in *E. coli*. These observations suggest that *yadB* might encode a protein that is important and peculiar to this class of bacteria.

Table 5.3.1 Genetic organisation flanking *pcnB* in *E. coli* and its homologues in other eubacterial species - 1.

Genetic organisation flanking *pcnB* in *E. coli* and its homologues in other eubacterial species. Organisms are tabulated according to the similarity of their *pcnB* homologue to *pcnB* in *E. coli* (most similar at the top). ✓, homologue present, as determined from the expected (E) values returned from a tblastn search with each gene from *E. coli*; ✘, no homologue present; ?, presence of gene not known; 1, completed genome; 2, sequences unassembled; Genes identified as homologues of *E. coli pcnB* positioned below the line may in fact be homologues of *E. coli cca* and possess tRNA nucleotidyltransferase activity. See text for details.

Organism	Homologues of *E. coli* genes						Genetic Organisation Flanking *pcnB* (5'-3')
	pcnB	*cca*	*folK*	*yadB*	*gltX*	*dksA*	
E. coli[1]	✓	✓	✓	✓	✓	✓	*dksA-yadB-pcnB-folK*
S. typhi	✓	✓	✓	✓	✓	✓	*dksA-yadB-pcnB-folK*
Y. pestis	✓	✓	✓	✓	✓	✓	*dksA-yadB-pcnB-folK*
H. influenzae[1]	✓	✓	✓	✘	✓	✓	*dksA-pcnB-folK*
P. aeruginosa	✓	✓	✓	✓	✓	✓	*dksA-yadB-pcnB-folK*
P. putida	✓	?	✓	?	?	?	*pcnB-folK*
B. pertussis	✓	✓	✓	✓	✓	✓	*pcnB-folK*; *yadB* & *dksA* on different contigs[2]
N. meningitidis serogroup A	✓	✓	✓	✓	✓	✓	*dksA* >35 kb from *gltX*; other genes on different contigs[2]
N. gonorrhoeae	✓	✓	✓	✓	✓	✓	All genes on different contigs[2]
A. actinomycetemcomitans	✓	✓	✓	?	✓	✓	*dksA-pcnB-folK*
V. cholerae	✓	✓	✓	✓	✓	✓	*pcnB-folK*; *yadB* & *dksA* on different contigs[2]
C. trachomatis[1]	✓	✓	✓	✘	✓	✓	*folK* >200 kb 3' of *pcnB*; *dksA* ~2.3 kb 5' of *pcnB*
T. pallidum[1]	✓	✓	✘	✘	✓	✓	*dksA* >500 kb 5' of *pcnB*
B. subtilis[1]	✓	✘	✓	✘	✓	✓	*folK* >2000 kb & *dksA* >1000 kb from *pcnB*
B. burgdorferi[1]	✓	✘	✘	✘	✓	✓	*dksA* >550 kb from *pcnB*; *rnc* homologue immediately 5' to *pcnB*
S. pneumoniae	✓	?	✓	?	✓	?	All genes on different contigs[2]
E. faecalis	✓	?	✓	?	✓	?	All genes on different contigs[2]
S. mutans	✓	?	✓	?	✓	?	All genes on different contigs[2]
S. pyogenes	✓	?	✓	?	✓	?	All genes on different contigs[2]
H. pylori[1]	✓	✘	✓	✓	✓	✘	*yadB* ~1 kb 3' of *pcnB*; *gltX* >185 kb & *folK* >405 kb from *pcnB*
A. aeolicus[1]	✓	✓	✓	✘	✓	✓	*folk* >180 kb & *dksA* >1200 kb from *pcnB*
M. tuberculosis[1] H37Rv	✘	✓	✓	✘	✓	✘	*folK* >350 kb & *gltX* >1000 kb from *cca*
R. prowazekii[1]	✓	✘	✘	✓	✓	✓	*dksA* >1000 kb from *pcnB*
C. jejuni	✘	✓	✓	✓	✓	✓	*folK* (79), *dksA* (127), *cca* (739), *yadB* (793), *gltX* (1221); numbers are coordinates in kb within the complete genome (1640 kb)
Synechocystis sp.[1]	✓	✓	✓	✘	✓	✘	*folK* >1500 kb & *dksA* >500 kb from *pcnB*
S. aureus	✓	?	✓	?	?	?	All genes on different contigs[2]

Table 5.3.2 Genetic organisation flanking *pcnB* in *E. coli* and its homologues in other eubacterial species - 2.

Genetic organisation flanking *pcnB* in *E. coli* and its homologues in other eubacterial species. Organisms are tabulated according to the similarity of their *pcnB* homologue to *pcnB* in *E. coli* (most similar at the top). Numbers are the expected (E) values returned from a tblastn search with each gene from *E. coli*; ✘, no homologue present; ?, presence of gene not known; 1, completed genome; 2, sequences unassembled; Genes identified as homologues of *E. coli pcnB* positioned below the line may in fact be homologues of *E. coli cca* and possess tRNA nucleotidyltransferase activity. See text for details.

Organism	Homologues of *E. coli* genes						Genetic Organisation Flanking *pcnB* (5'-3')
	pcnB	*cca*	*folK*	*yadB*	*gltX*	*dksA*	
E. coli[1]	0.0	0.0	1e-89	0.0	0.0	5e-84	*dksA-yadB-pcnB-folK*
S. typhi	0.0	e-158	1e-72	e-145	2e-47	7e-12	*dksA-yadB-pcnB-folK*
Y. pestis	0.0	e-180	2e-53	e-108	0.0	9e-80	*dksA-yadB-pcnB-folK*
H. influenzae[1]	e-135	e-139	3e-48	✘	0.0	6e-60	*dksA-pcnB-folK*
P. aeruginosa	e-114	e-139	1e-40	1e-73	1e-80	2e-55	*dksA-yadB-pcnB-folK*
P. putida	e-108	?	8e-26	?	?	?	*pcnB-folK*
B. pertussis	9e-93	8e-70	1e-21	1e-53	e-132	2e-23	*pcnB-folK*; *yadB* & *dksA* on different contigs[2]
N. meningitidis serogroup A	1e-82	e-117	2e-21	2e-56	e-121	2e-26	*dksA* >35 kb from *gltX*; other genes on different contigs[2]
N. gonorrhoeae	4e-80	e-119	4e-21	6e-57	e-119	2e-26	All genes on different contigs[2]
A. actinomycetemcomitans	4e-72	1e-94	1e-44	?	0.0	1e-61	*dksA-pcnB-folK*
V. cholerae	5e-67	2e-42	3e-50	1e-62	e-100	2e-69	*pcnB-folK*; *yadB* & *dksA* on different contigs[2]
C. trachomatis[1]	4e-58	2e-16	3e-10	✘	1e-64	1e-04	*folK* >200 kb 3' of *pcnB*; *dksA* ~2.3 kb 5' of *pcnB*
T. pallidum[1]	1e-43	8e-17	✘	✘	1e-80	4e-04	*dksA* >500 kb 5' of *pcnB*
B. subtilis[1]	2e-33	✘	1e-25	✘	9e-95	7e-05	*folK* >2000 kb & *dksA* >1000 kb from *pcnB*
B. burgdorferi[1]	6e-28	✘	✘	✘	2e-83	5e-04	*dksA* >550 kb from *pcnB*; *rnc* homologue immediately 5' to *pcnB*
S. pneumoniae	8e-26	?	1e-13	?	2e-54	?	All genes on different contigs[2]
E. faecalis	1e-25	?	5e-22	?	2e-93	?	All genes on different contigs[2]
S. mutans	2e-24	?	3e-26	?	4e-32	?	All genes on different contigs[2]
S. pyogenes	4e-23	?	6e-29	?	1e-87	?	All genes on different contigs[2]
H. pylori[1]	9e-21	✘	3e-11	3e-41	1e-98	✘	*yadB* ~1 kb 3' of *pcnB*; *gltX* >185 kb & *folK* >405 kb from *pcnB*
A. aeolicus[1]	2e-19	9e-18	1e-20	✘	e-102	1e-14	*folk* >180 kb & *dksA* >1200 kb from *pcnB*
M. tuberculosis[1] H37Rv	✘	2e-22	2e-11	✘	1e-71	✘	*folK* >350 kb & *gltX* >1000 kb from *cca*
R. prowazekii[1]	4e-16	✘	✘	3e-37	3e-88	5e-21	*dksA* >1000 kb from *pcnB*
C. jejuni	✘	4e-20	2e-08	2e-37	e-106	1e-07	*folK* (79), *dksA* (127), *cca* (739), *yadB* (793), *gltX* (1221); numbers are coordinates in kb within the complete genome (1640 kb)
Synechocystis sp.[1]	1e-11	2e-12	1e-24	✘	3e-92	✘	*folK* >1500 kb & *dksA* >500 kb from *pcnB*
S. aureus	3e-10	?	9e-17	?	?	?	All genes on different contigs[2]

Table 5.3.3 Taxonomic classification of eubacterial species that possess a homologue of *E. coli pcnB*.

Taxonomic classification of eubacterial species that possess a homologue of *E. coli pcnB*. Organisms are tabulated according to the similarity of their *pcnB* homologue to *pcnB* in *E. coli* (most similar at the top). GP, Gram-positive (or Firmicutes); N/A, not applicable. The proteobacteria (or purple non-sulphur bacteria) contain most species traditionally thought of as gram-negative eubacteria; the Chlamydiae and Spirochaetes are also considered to be gram-negative; low G+C gram-positive bacteria are also known as the Bacillus/Clostridium group; high G+C gram-positive bacteria are also known as Actinomycetes (or Actinobacteria); Cyanobacteria are also known as blue-green algae.

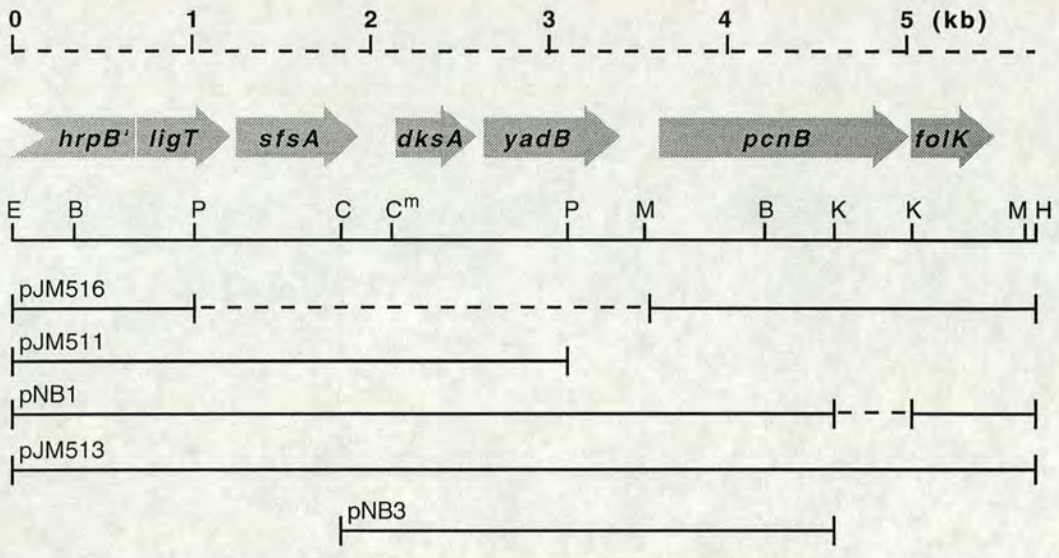
Organism	Class	Subdivision	Order	Family	Genome Sequence
Escherichia coli	Proteobacteria	γ		Enterobacteriaceae	Completed
Salmonella typhi	Proteobacteria	γ		Enterobacteriaceae	Unfinished
Yersinia pestis	Proteobacteria	γ		Enterobacteriaceae	Unfinished
Haemophilus influenzae	Proteobacteria	γ		Pasteurellaceae	Completed
Pseudomonas aeruginosa	Proteobacteria	γ		Pseudomonas gp.	Unfinished
Pseudomonas putida	Proteobacteria	γ		Pseudomonas gp.	Unfinished
Bordetella pertussis	Proteobacteria	β		Alcaligenaceae	Unfinished
Neisseria meningitidis serogroup A	Proteobacteria	β		Neisseriaceae	Unfinished
Neisseria gonorrhoeae	Proteobacteria	β		Neisseriaceae	Unfinished
Actinobacillus actinomycetemcomitans	Proteobacteria	γ		Pasteurellaceae	Unfinished
Vibrio cholerae	Proteobacteria	γ		Vibrionaceae	Unfinished
Chlamydia trachomatis	Chlamydiae	N/A	Chlamydiales	Chlamydiaceae	Completed
Treponema pallidum	Spirochaetes	N/A	Spirochaetales	Spirochaetaceae	Completed
Bacillus subtilis	GP; low G+C	N/A		Bacillaceae	Completed
Borrelia burgdorferi	Spirochaetes	N/A	Spirochaetales	Spirochaetaceae	Completed
Streptococcus pneumoniae	GP; low G+C	N/A		Streptococcaceae	Unfinished
Enterococcus faecalis	GP; low G+C	N/A		Enterococcaceae	Unfinished
Streptococcus mutans	GP; low G+C	N/A		Streptococcaceae	Unfinished
Streptococcus pyogenes	GP; low G+C	N/A		Streptococcaceae	Unfinished
Helicobacter pylori	Proteobacteria	ε		Helicobacter gp.	Completed
Aquifex aeolicus		N/A	Aquificales	Aquificaceae	Completed
Mycobacterium tuberculosis H37Rv	GP; high G+C	N/A	Actinomycetales	Mycobacteriaceae	Completed
Rickettsia prowazekii	Proteobacteria	α	Rickettsiales	Rickettsiaceae	Completed
Campylobacter jejuni	Proteobacteria	ε		Campylobacter gp.	Unfinished
Synechocystis sp.	Cyanobacteria	N/A	Chroococcales		Completed
Staphylococcus aureus	GP; low G+C	N/A		Bacillaceae	Unfinished

5.4 Expression of the *yadB* gene product

5.4.1 Cell-free coupled transcription/translation of *yadB*

Assuming that *yadB* is a real gene, an attempt was made to visualise its protein product by SDS-PAGE. *In vitro* DNA-directed protein translation was carried out from pNB1 in addition to several existing plasmid templates (Figures 5.4.1 and 5.4.2). [^{3}H] leucine was used as the radiolabel in order to enhance the signal from YadB, as the sequence data for this predicted gene-product indicated the presence of only one internal methionine residue. A band of approximately 17 kDa translated from pJM511 was the only possible candidate YadB protein. Although the predicted size of YadB (based on current sequence data from Blattner *et al*, 1997) is 34.7 kDa (308 amino acids), the pJM511 insert bears the DNA for just the N-terminal half of the open reading frame (155 amino acids). The size of this truncated protein would be 17.8 kDa (Figure 5.4.1). If the 17 kDa protein identified in Figure 5.4.2 is a YadB′ deletion derivative, the full-length protein should be visible in pJM513 and pNB1. However, there is no visible product from the reaction using these two plasmids that correlates with the predicted size of full-length YadB. It is possible that co-migration with the *bla* gene-product (31.5 kDa) might mask the presence of full-length YadB, but a comparison of the band densitiy for β-lactamase from pJM516 (which does not encode the *yadB* open reading frame) with the equivalent sized band from pJM513 (which does encode *yadB*) shows no evidence of two superimposed proteins. The *Hin*d III boundary between insert and vector unique to pJM511 could be the source of an encoded fusion protein that migrates as a 17 kDa band. However, inspection of the nucleotide sequence spanning the region found no evidence to support this. Although *folK* encodes a polypeptide with a compositional molecular mass of 17.9 kDa, the unidentified 17 kDa band cannot be its gene product because the pJM511 insert does not include its nucleotide sequence (Figure 5.4.1). Indeed FolK is not even visible in the *in vitro* translation products from templates that do include its open reading frame (pJM513 and pJM516; in pNB1, *pcnB* minus its carboxy terminus is fused to *folK*), presumably because it is expressed at a very low level. Some enzymes of the folate pathway are known to be present in only minute quantities; FolK has been shown to compose less than 0.01% of the total soluble cell proteins (Talarico *et al*, 1991).

Figure 5.4.1 pJM513 and its deletion derivatives.



pJM513 and its deletion derivatives. Bars below the line represent the extent of DNA present in each insert. Hyphenated lines represent the region of deleted DNA. Arrows on coding regions indicate the direction of transcription. Relevant restriction sites: B, *Bst*E II; C, *Cla* I; C^m, *Cla* I blocked by overlapping *dam* methylation; E, *Eco*R I; H, *Hin*d III; K, *Kpn* I; M, *Mfe* I; P, *Pvu* II. Map is drawn to scale.

5.4.2 *In vivo* expression of the *yadB* gene product - 1

The 17 kDa band expressed *in vitro* from pJM511 was considered unlikely to be the product of a *yadB* truncation mutant. In order to eliminate the possibility of YadB being masked by other gene-products with similar migration rates, the DNA encoding *yadB* was subjected to *in vivo* overexpression from a new construct (pNB3) that only encodes one flanking gene in full, other than *yadB*. The 2.7 kb *Cla* I/*Kpn* I fragment from pNB1 was cloned into the *Acc* I/*Kpn* I sites in the polylinker of pUC19 creating pNB2. This step was carried out in order to acquire flanking restriction sites that would facilitate cloning the fragment into the target vector (note: two *Cla* I sites exist in the pNB1 insert; one is blocked by overlapping *dam* methylation). The 2.8 kb *Eco*R I/*Hin*d III fragment from pNB2 was cloned into the corresponding sites in pT7-4 forming pNB3 (Figure 5.4.3; note: as the two fragments resulting from this digestion were expected to be very similar in size, *Sca* I was added to the reaction in order to cut the pUC19 backbone in two).

Figure 5.4.2 *In vitro* translation of the genes encoded by pJM513 and its deletion derivatives.

In vitro translation products were labelled with [^{3}H] leucine and separated on a 12% SDS-polyacrylamide gel. The deletion of the 419 bp *Kpn* I fragment from pNB1 creates a predicted PcnB'-FolK fusion protein of 55.6 kDa (cf. 53.9 kDa for PcnB seen in pJM513 and pJM516). The 9 kDa protein (labelled Cmp fusion) is believed to be a chloramphenicol acetyl transferase fusion protein encoded at the boundary between insert and vector (March *et al*, 1989). Bla, β-lactamase, is encoded by the ampicillin resistance gene; Cmp, Chloramphenicol acetyl transferase, is encoded by the chloramphenicol resistance gene; other proteins are identified in Figure 5.4.1; pBR328 is the parental vector of pJM513 and its derivatives.

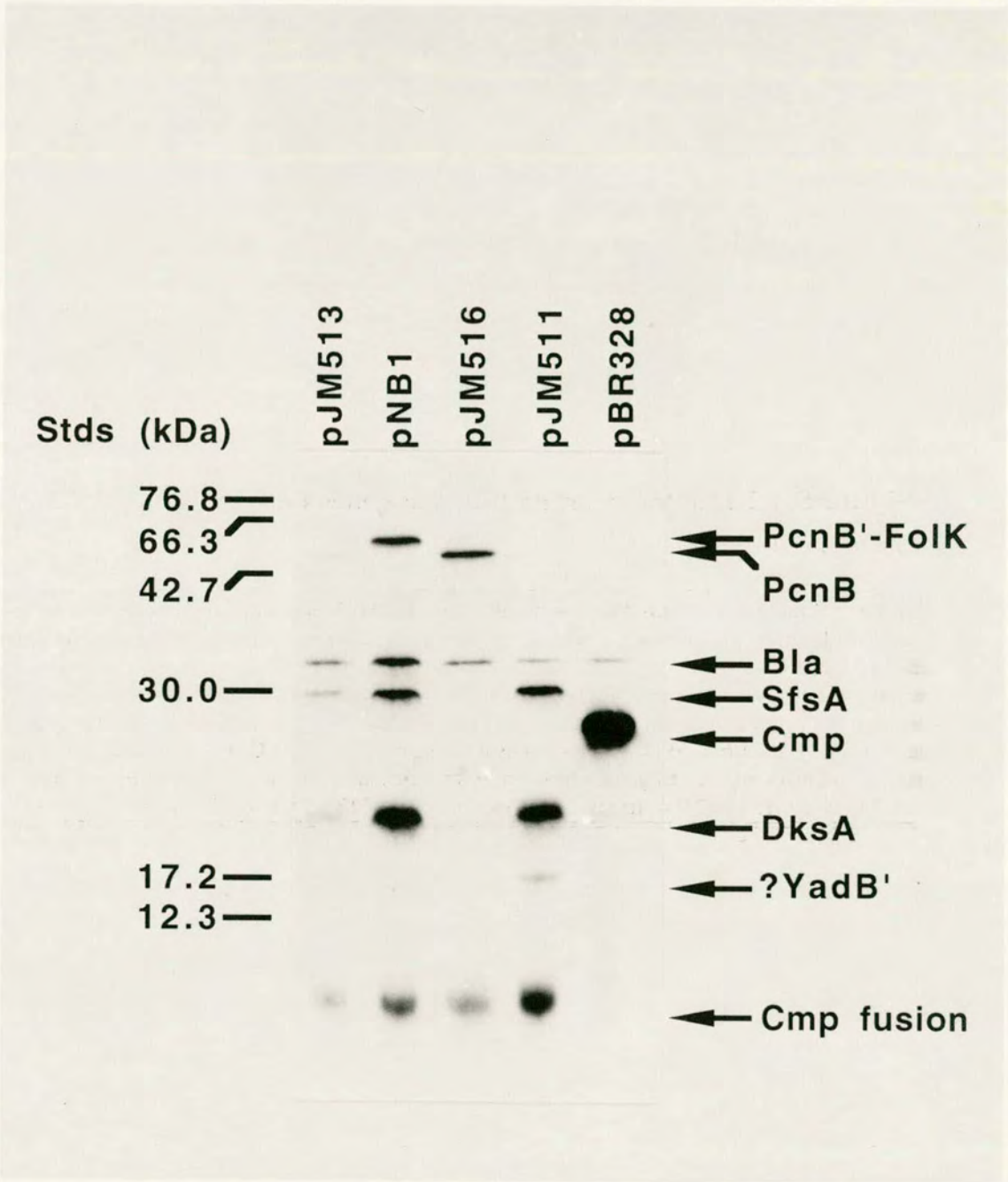
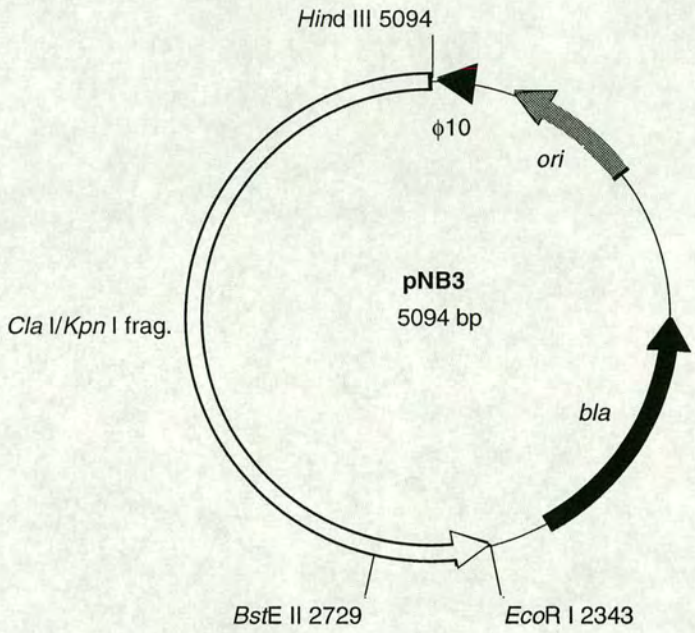

Stds (kDa)
pJM513
pNB1
pJM516
pJM511
pBR328
76.8
66.3
42.7
30.0
17.2
12.3
PcnB'-FolK
PcnB
Bla
SfsA
Cmp
DksA
?YadB'
Cmp fusion

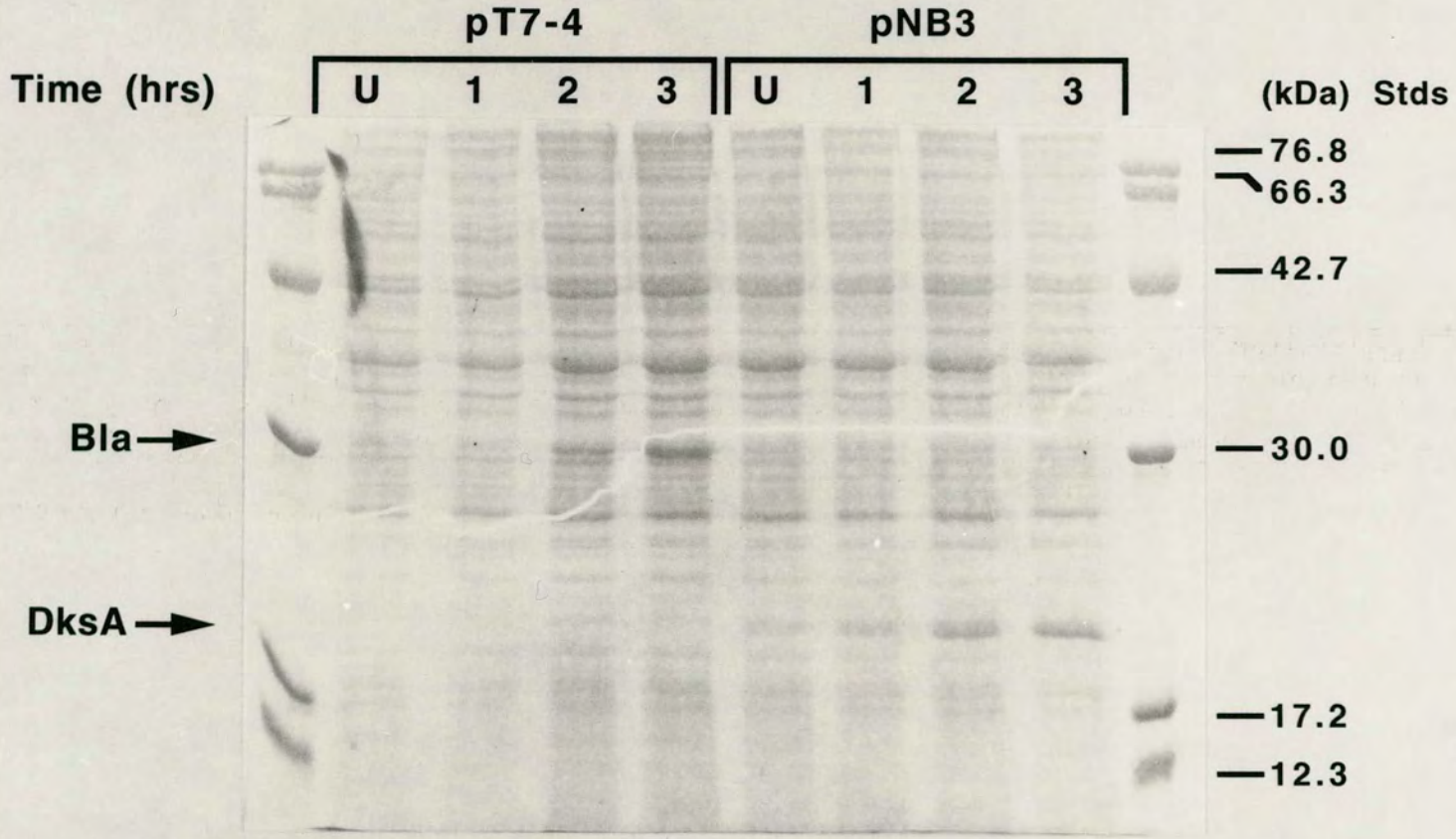
Figure 5.4.3 pNB3.



Any gene cloned downstream from the T7 RNA polymerase promoter (ϕ10) in pT7-4 will be overexpressed *in vivo* when transformed into a strain lysogenised with a phage λ encoding T7 RNA polymerase (section 2.7.3). BL21(DE3)/pLysS harbouring either pNB3 or its parental vector, pT7-4, was treated with IPTG to induce expression of T7 RNA polymerase and hence plasmid-encoded genes, for up to three hours (Figure 5.4.4). An approximately 18 kDa protein appears to be the only product of overexpression from pNB3. Based on sequence data, the results from *in vitro* translation (Figure 5.4.2) and the region of DNA known to be present in pNB3, it is surmised that this protein is DksA.

Figure 5.4.4 Overexpression of Proteins *In Vivo* from pNB3.

Overexpression of proteins *in vivo* from pNB3. Cultures of BL21(DE3)/pLysS harbouring the indicated plasmids were grown in Spizizen minimal medium at 37°C to OD_{600} of approximately 0.8. Expression was induced with 0.5 mM IPTG for up to three hours. Proteins were separated on a 12% SDS-polyacrylamide gel. Time (hrs), indicates the post-induction incubation interval in hours. U, is an uninduced control. Overexpression of β-lactamase (Bla) can clearly be seen from pT7-4, the parental vector control. An 18 kDa protein believed to be DksA is expressed from pNB3.

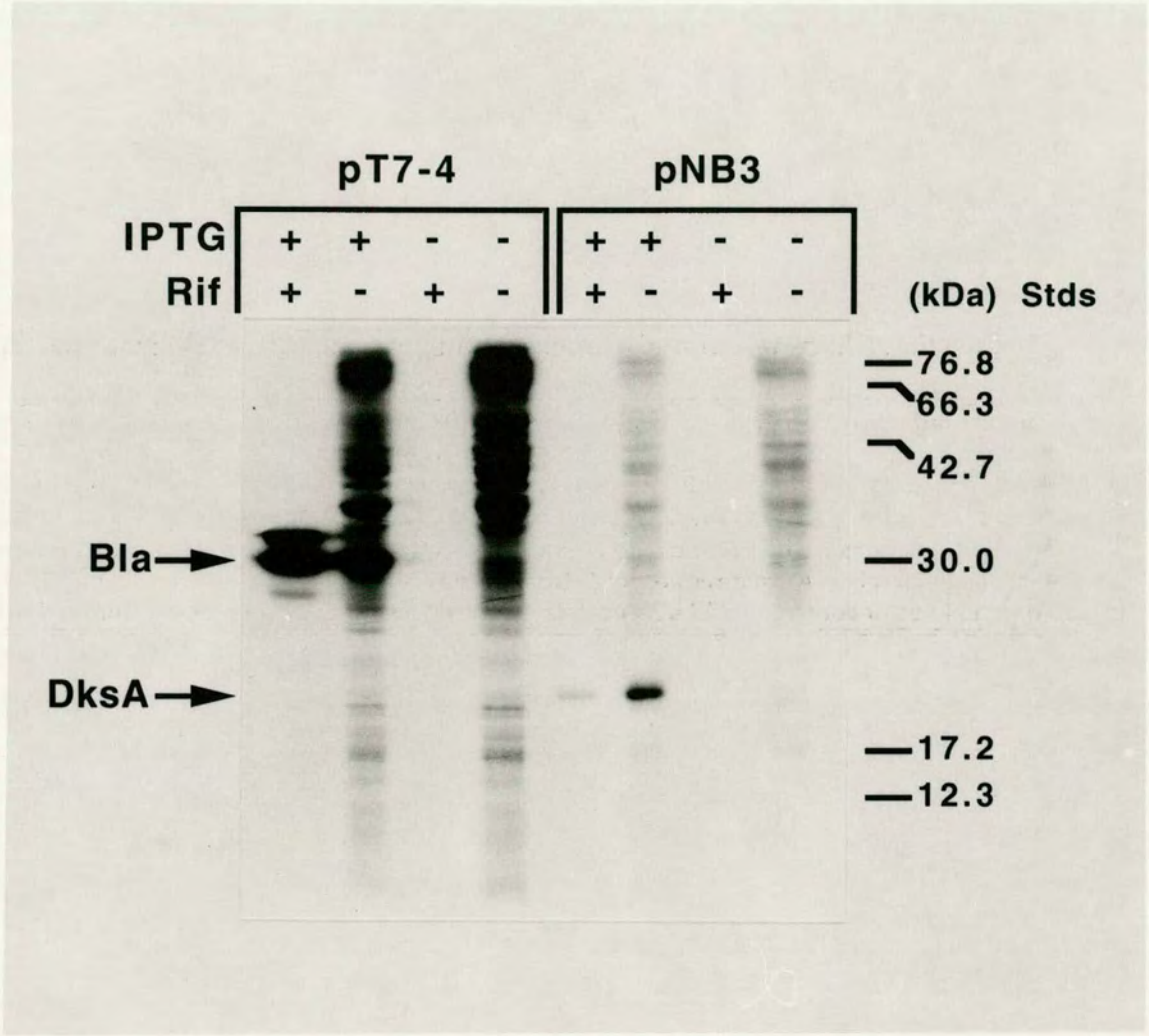

pT7-4
pNB3
Time (hrs)
U
1
2
3
U
1
2
3
(kDa) Stds
76.8
66.3
42.7
Bla
30.0
DksA
17.2
12.3

5.4.3 *In vivo* expression of the *yadB* gene product - 2

The system employed to detect YadB protein from pNB3 was not very selective (Section 5.4.2). By incorporating a radiolabel in the presence of rifampicin, an antibiotic that blocks transcription initiation from native promoters by inhibiting the activity of *E. coli* RNA polymerase, it was possible to increase the specificity and sensitivity of the procedure by eliminating background host expression. However, in spite of the absence of interfering expression from the host, DksA was still the only visible pNB3-encoded protein (Figure 5.4.5). This was unexpected. Not because a candidate YadB remained undetected, but rather, because the expression of β-lactamase from the *bla* gene downstream from the cloned DNA was not visible (also the case for *pcnB'*, which is present). The transcriptional orientation of the *bla* gene in pT7-4 is the same as the ϕ10 promoter for T7 RNA polymerase. Expression of β-lactamase directed by this enzyme from the parental vector, pT7-4, is clearly visible. It would therefore appear that the T7 RNA polymerase-directed transcription of pNB3 is subject to premature termination. A T7 RNA polymerase-sensitive termination sequence, located shortly after the 3'-end of *dksA*, could exert an inhibitory polar effect on the expression of *yadB* and *bla*, located downstream from the termination element. The lack of T7 RNA polymerase-directed transcription of *bla* could also be due to the higher elongation speed (compared with *E. coli* RNA polymerase) characteristic of this enzyme. In prokaryotes, transcription and translation are usually synchronous, so that no extensive gap exists between the transcribing RNA polymerase and the leading ribosome. It has been proposed that the eight-fold faster T7-derived enzyme could, in some cases, leave the nascent mRNA exposed and susceptible to endonucleolytic cleavage (and hence degradation) ahead of the leading ribosome. This may seem an unlikely explanation, as pT7-4 should be similarly affected. However, an endonuclease-sensitive site (such as an RNase E or RNase III target sequence) might have been introduced into pNB3 on the cloned fragment, which is unmasked by the high elongation speed of the T7 RNA polymerase (Lopez *et al*, 1994).

Figure 5.4.5 *In vivo* Radiolabelling of Proteins Overexpressed from pNB3.

In vivo radiolabelling of proteins overexpressed from pNB3. Cultures of BL21(DE3)/pLysS harbouring the indicated plasmids were grown in Spizizen minimal medium at 37°C to OD_{600} of approximately 0.8. Expression was induced with 0.5 mM IPTG for 80 minutes. *In vivo* translation products were labelled with [^{3}H] leucine and separated on a 12% SDS-polyacrylamide gel. The presence or absence of rifampicin (Rif) and IPTG is indicated by a + or - respectively. Overexpression of β-lactamase (Bla) can clearly be seen from pT7-4, the parental vector control. An 18 kDa protein believed to be DksA is expressed from pNB3.


pT7-4
pNB3
IPTG
Rif
+ + - - + + - -
+ - + - + - + -
(kDa) Stds
76.8
66.3
42.7
30.0
17.2
12.3
Bla
DksA

5.5 Summary

Four homologues of *E. coli pcnB* are identified as potentially able to utilise an AUU or AUC triplet as a translational initiation codon. The organisms they come from all belong to the γ subdivison of the Proteobacteria: two from the family Enterobacteriaceae, and two from the family Pasteurellaceae. Recent work by others in this laboratory supports the findings of the sequence analysis described in this chapter: when interrupted with an amber mutation at an appropriate position within the open reading frame, the *pcnB* homologue from *H. influenzae* is unable to complement a Δ*pcnB* strain. A run of T residues is present in each of these homologues, located exactly nine nucleotides downstream from the -10 element of a σ^{70} promoter (identified by examining the sequence upstream from the potential ATT initiation codon). The presence of this run of Ts in each gene is assumed to be important. Possible functions include the destabilising poly(T) tract of a rho-independent transcriptional terminator, or a control element that confers UTP-dependent reiterative transcription as a mechanism for regulating the expression of *pcnB*. A number of eubacteria were found to share with *E. coli* the equivalent genetic organisation flanking their homologue of *pcnB*. This database search was carried out as a first step toward establishing a functional relationship between *pcnB* and *folK*, and to determine whether or not *yadB* is a real gene. The *dksA-yadB-pcnB-folK* organisation in *E. coli* was only found in organisms belonging to the γ subdivision of the Proteobacteria. The *pcnB-folK* organisation is also preserved in *B. pertussis*, a more distantly related organism belonging to the β subdivision. A *yadB* homologue is present in representatives from almost all of the subdivisions of the Proteobacteria, with the striking exception of *H. influenzae*. However, it is only found located between *dksA* and *pcnB* in bacteria belonging to the γ subdivision. The wide distribution of *yadB* within the Proteobacteria suggests this open reading frame has a functional role. *In vitro* and *in vivo* translation from several templates failed to visualise the expression of the *yadB* gene product. As this approach was not exhaustive, the apparent absence of a *yadB*-associated polypeptide cannot be considered conclusive evidence that the open reading frame is not expressed.

CHAPTER 6

References

References

Adhya, S. (1987) The Galactose Operon. In: Neidhardt, F.C., ed. *Escherichia coli and Salmonella: Cellular and Molecular Biology, 2*. Washington, DC: American Society for Microbiology. pp 1503-1512.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25:** 3389-3402.

Arn, E.A. and Abelson, J.N. (1996) The 2'-5' RNA ligase of *Escherichia coli* - Purification, cloning, and genomic disruption. *Journal of Biological Chemistry* **271:** 31145-31153.

August, J.T., Oritz, P.J. and Hurwitz, J. (1962) Ribonucleic acid-dependent ribonucleotide incorporation: I. Purification and properties of the enzyme. *Journal of Biological Chemistry* **237:** 3786-3793.

Backman, K., Betlach, M., Boyer, H.W. and Yanofsky, S. (1978) Genetic and physical studies on the replication of ColE1-type plasmids. *Cold Spring Harbor Symposia on Quantitative Biology* **43:** 69-76.

Bardwell, J.C.A., Regnier, P., Chen, S.M., Nakamura, Y., Grunberg-Manago M. and Court D.L. (1989) Auto-regulation of RNase III operon by messenger RNA processing. *EMBO Journal* **8:** 3401-3407.

Becker, G. and Hengge-Aronis, R. (2001) What makes an *Escherichia coli* promoter σ^S dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of σ^S. *Molecular Microbiology* **39:** 1153-1165.

Berkhout, B., Vanderlaken, C.J. and van Knippenberg, P.H. (1986) Formylmethionyl transfer RNA-binding to 30S ribosomes programmed with homopolynucleotides and the effect of translational initiation factor-III. *Biochimica et Biophysica Acta* **866:** 144-153.

Berlyn, M.K.B. (1998) Linkage map of Escherichia coli K-12, edition 10: The traditional map. *Microbiology and Molecular Biology Reviews* **62:** 814-984.

Binnie, U., Wong, K., McAteer, S. and Masters, M. (1999) Absence of RNase III alters the pathway by which RNA I, the antisense inhibitor of ColE1 replication, decays. *Microbiology-UK* **145:** 3089-3100.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7:** 1513-1523.

Blair, D.G. and Helinski, D.R. (1975) Relaxation complexes of plasmid DNA and protein. *Journal of Biological Chemistry* **250:** 8785-8789 and following papers.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453-1462.

Blomberg, P., Wagner, E.G.H. and Nordstrom, K. (1990) Control of replication of plasmid R1 - the duplex between the antisense RNA, *copA*, and its target, *copT*, is processed specifically *in vivo* and *in vitro* by RNase III. EMBO Journal **9:** 2331-2340.

Bolivar, F. (1978) Construction and characterisation of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique *Eco*R I sites for selection of *Eco*R I-generated recombinant DNA molecules. *Gene* **4:** 121-136.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2:** 95-113.

Boni, I.V., Isaeva, D. M., Musychenko, M.L. and Tzareva, N.V. (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Research* **19:** 155-162.

Bouvet, P. and Belasco, J.G. (1992) Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in *Escherichia coli*. *Nature* **360:** 488-491.

Brenner, M. and Tomizawa, J-I. (1991) Quantitation of ColE1-encoded replication elements. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 405-409.

Breton, R., Sanfacon, H., Papayannopoulos, I., Biemann, K. and La Pointe, J. (1986) Glutamyl-transfer RNA-synthetase of *Escherichia coli*: isolation and primary structure of the *gltx* gene and homology with other aminoacyl-transfer RNA-synthetases. *Journal of Biological Chemistry* **261:** 610-617.

Brombach, M. and Pon. C.L. (1987) The unusual translational initiation codon AUU limits the expression of the *infC* (initiation-factor IF3) gene of *Escherichia coli*. *Molecular & General Genetics* **208:** 94-100.

Butler, J.S, Springer, M., Dondon, J., Graffe, M. and Grunberg-Manago, M. (1986)*Escherichia coli* protein-synthesis initiation-factor IF3 controls its own gene-expression at the translational level *in vivo*. *Journal of Molecular Biology* **192:** 767-780.

Butler, J.S., Springer, M. and Grunberg-Manago, M. (1987) AUU-to-AUG mutation in the initiator codon of the translation initiation-factor IF3 abolishes translational autocontrol of its own gene (*infC*) *in vivo*. *Proceedings of The National Academy of Sciences of The United States of America* **84:** 4022-4025.

Bycroft, M., Hubbard, T.J.P., Proctor, M., Freund, S.M.V. and Murzin, A.G. (1997) The solution structure of the S1 RNA binding domain: A member of an ancient nucleic acid-binding fold. *Cell* **88:** 235-242.

Cao, G.J., Pogliano, J. and Sarkar, N. (1996) Identification of the coding region for a second poly(A) polymerase in *Escherichia coli*. *Proceedings of The National Academy of Sciences of The United States of America* **93:** 11580-11585.

Cao, G.J. and Sarkar, N. (1992) Identification of the gene for an *Escherichia coli* poly(A) polymerase. *Proceedings of The National Academy of Sciences of The United States of America* **89:** 10380-10384.

Carpousis, A.J. and Gralla, J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the *lac* UV5 promoter. *Biochemistry* **19:** 3245-3253.

Carpousis, A.J., van Houwe, G., Ehretsmann, C. and Krisch, H.M. (1994) Copurification of *Escherichia coli* RNase-E and PNPase - evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* **76:** 889-900.

Carpousis, A.J., Vanzo, N.F. and Raynal, L.C. (1999) mRNA degradation - a tale of poly(A) and mutiprotein machines. *Trends in Genetics* **15:** 24-28.

Cesareni, G., Helmer-Ciierich, M. and Castagnoli, L. (1991) Control of ColE1 plasmid replication by antisense RNA. *Trends in Genetics* **7:** 230-235.

Cesareni, G., Muesing, M.A. and Polisky, B. (1982) Control of ColE1 replication: the *rop* gene product negatively affects transcription from the replication primer promoter. *Proceedings of the National Academy of Sciences of the United States of America* **79:** 6313-6317.

Chang, A.C.Y. and Cohen, S.N. (1978) Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. *Journal of Bacteriology* **134:** 1141-1156.

Chattoraj, D.K. (2000) Control of plasmid DNA replication by iterons: no longer paradoxical. *Molecular Microbiology* **37:** 467-476.

Cheng, Y.L., Kalman, L.V. and Kaiser, D. (1994) The *dsg* gene of *Myxococcus xanthus* encodes a protein similar to translation initiation-factor IF3. *Journal of Bacteriology* **176:** 1427-1433.

Cheng, Z.F., Zuo, Y.H., Li, Z.W., Rudd, K.E., and Deutscher, M.P. (1998). The *vacB* gene required for virulence in *Shigella flexneri* and *Escherichia coli* encodes the exoribonuclease RNase R. *Journal of Biological Chemistry* **273:** 14077-14080.

Churchward, G., Belin, D. and Nagamine, Y. (1984) A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene* **31:** 165-171.

Clewell, D.B. (1972) Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. *Journal of Bacteriology* **110:** 667-676.

Coburn, G.A. and Mackie, G.A. (1996a) Overexpression, purification, and properties of *Escherichia coli* ribonuclease II. *Journal of Biological Chemistry* **271:** 1048-1053.

Coburn, G.A. and Mackie, G.A. (1996b) Differential sensitivities of portions of the mRNA for ribosomal protein S20 to 3'-exonucleases dependent on oligoadenylation and RNA secondary structure. *Journal of Biological Chemistry* **271:** 15776-15781.

Coburn, G.A. and Mackie, G.A. (1999) Degradation of mRNA in *Escherichia coli*: An old problem with some new twists. *Progress in Nucleic Acid Research and Molecular Biology* **62:** 55-108.

Coburn, G.A., Miao, X., Briant, D.J. and Mackie, G.A. (1999) Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. *Genes & Development* **13:** 2594-2603.

Cohen, S.N. (1993) Bacterial plasmids: their extraordinary contribution to molecular genetics. *Gene* **135:** 67-76.

Cole, J.R., Olsson, C.L., Hershey, J.W.B., Grunberg-Manago, M. and Nomura, M. (1987) Feedback-regulation of ribosomal-RNA synthesis in *Escherichia coli* - requirement for initiation-factor IF2 *Journal of Molecular Biology* **198:** 383-392.

Cudny, H., Lupski, J.R., Godson, G.N. and Deutscher, M.P. (1986) Cloning, sequencing, and species relatedness of the *Escherichia coli cca* gene encoding the enzyme transfer-RNA nucleotidyltransferase. *Journal of Biological Chemistry* **261:** 6444-6449.

de Cock, E., Springer, M. and Dardel, F. (1999) The interdomain linker of *Escherichia coli* initiation factor IF3: a possible trigger of translation initiation specificity. *Molecular Microbiology* **32:** 193-202.

del Solar, G. and Espinosa, M. (2000) Plasmid copy number control: an ever-growing story. *Molecular Microbiology* **37:** 492-500.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M. and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews* **62:** 434-464.

de Smit, M.H. and van Duin, J. (1990a) Secondary structure of the ribosome binding-site determines translational efficiency - a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America* **87:** 7668-7672.

de Smit, M.H. and van Duin, J. (1990b) Control of prokaryotic translational initiation by messenger-RNA secondary structure. *Progress in Nucleic Acid Research and Molecular Biology* **38:** 1-35.

de Smit, M.H. and van Duin, J. (1994) Translational initiation on structured messengers - another role for the Shine-Dalgarno interaction. *Journal of Molecular Biology* **235:** 173-184.

Deutscher, M.P. (1988) The metabolic role of RNases. *Trends in Biochemical Sciences* **13:** 136-139.

Deutscher, M.P. (1993a) Ribonuclease multiplicity, diversity, and complexity. *Journal of Biological Chemistry* **268:** 13011-13014.

Deutscher, M.P. (1993b) Promiscuous exoribonucleases of *Escherichia coli*. *Journal of Bacteriology* **175:** 4577-4583.

Donovan, W.P. and Kushner, S.R. (1986). Polynucleotide phosphorylase and ribonuclease-II are required for cell viability and messenger-RNA turnover in *Escherichia coli* K-12. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **83:** 120-124.

Dunn, J.J. (1976) RNase III cleavage of single-stranded RNA. Effect of ionic strength on the fidelity of cleavage. *Journal of Biological Chemistry* **251:** 3807-3814.

Eguchi, Y., Itoh, T. and Tomizawa, J-I. (1991) Antisense RNA. *Annual Review of Biochemistry* **60:** 631-652.

Ellinger, T., Behnke, D., Bujard, H. and Gralla, J.D. (1994) Stalling of *Escherichia coli* RNA polymerase in the +6 to +12 region *in vivo* is associated with tight-binding to consensus promoter elements. *Journal of Molecular Biology* **239:** 455-465.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* RD. *Science* **269:** 496-512.

Fujita, N., Mori, H., Yura, T. and Ishihama, A. (1994) Systematic sequencing of the *Escherichia coli* genome - analysis of the 2.4-4.1 min (110,917-193,643 bp) region. *Nucleic Acids Research* **22:** 1637-1639.

Gausing, K. (1980) Regulation of Ribosome biosynthesis in *Escherichia coli.* In: Chambliss, G., Craven, G.R., Davies, J., Davis, K., Kahan, L. and Nomura, M., ed. *Ribosomes: structure, function and genetics*. Baltimore: University Park Press. pp 693-718.

Genetics Computer Group (GCG), Wisconsin Package Version 10.0 (1999), 575 Science Drive, Madison, Wisconsin, USA 53711.

Gibbs, T.W., Gill, D.R. and Salmond, G.P.C. (1992) Localized mutagenesis of the *fts YEX* operon - conditionally lethal missense substitutions in the FtsE cell-division protein of *Escherichia coli* are similar to those found in the cystic-fibrosis transmembrane conductance regulator protein (CFTR) of human patients. *Molecular & General Genetics* **234:** 121-128.

Gibson, T.J. (1984) Studies on the Epstein-Barr virus genome. Ph.D thesis, Cambridge University, England.

Ghosh, S. and Deutscher, M.P. (1999) Oligoribonuclease is an essential component of the mRNA decay pathway. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 4372-4377.

Gold, L. (1988) Posttranscriptional regulatory mechanisms in *Escherichia coli. Annual Review of Biochemistry* **57:** 199-233.

Gold, L., Stormo, G. and Saunders, R. (1984) *Escherichia coli* translational initiation factor-IF3 - a unique case of translational regulation. *Proceedings of The National Academy of Sciences of The United States of America-Biological Sciences* **81:** 7061-7065.

Gualerzi, C.O. and Pon, C.L. (1990) Initiation of messenger-RNA translation in prokaryotes. *Biochemistry* **29:** 5881-5889.

Gupta, R.S., Kasai, T. and Schlessinger, D. (1977) Purification and some novel properties of *Escherichia coli* RNase II . *Journal of Biological Chemistry* **252:** 8945-8949.

Hajnsdorf, E., Steier, O., Coscoy, L., Teysset, L. and Regnier, P. (1994a) Roles of RNase-E, RNase-II and PNPase in the degradation of the *rpsO* transcripts of *Escherichia coli* - stabilizing function of RNase-II and evidence for efficient degradation in an *ams-pnp-rnb* mutant. *EMBO Journal* **13:** 3368-3377.

Hajnsdorf, E., Carpousis, A.J. and Regnier, P. (1994b) Nucleolytic inactivation and degradation of the RNase III processed *pnp* message encoding polynucleotide phosphorylase of *Escherichia coli*. *Journal of Molecular Biology* **239:** 439-454.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166:** 557-580.

Hartz, D., Binkley, J., Hollingsworth, T. and Gold, L. (1990) Domains of initiator transfer-RNA and initiation codon crucial for initiator transfer-RNA selection by *Escherichia coli* IF3. *Genes & Development* **4:** 1790-1800.

Hartz, D., McPheeters, D.S., Traut, R. and Gold, L. (1988) Extension inhibition analysis of translation initiation-complexes. *Methods in Enzymology* **164:** 419-425.

He, L., Soderbom, F., Wagner, E.G.H., Binnie, U., Binns, N. and Masters, M. (1993) PcnB is required for the rapid degradation of RNAI, the antisense RNA that controls the copy number of ColE1-related plasmids. *Molecular Microbiology* **9:** 1131-1142.

Helinski, D.R., Toukdarian, A.E. and Novick, R.P. (1996) Replication control and other stable maintenance mechanisms of plasmids. In: Neidhardt, F.C., ed. *Escherichia coli and Salmonella: Cellular and Molecular Biology, 2*. Washington, DC: American Society for Microbiology. pp 2295-2324.

Hong, J.S. and Ames, B.N. (1971) Localised mutagenesis of any small region of the bacterial chromosome. *Proceedings of The National Academy of Sciences of The United States of America-Biological Sciences* **68:** 3158-3161.

Horecker, B.L., Thomas, J. and Monod, J. (1960) Galactose transport in *Escherichia coli*. I. General properties as studied with a galactokinaseless mutant. *Journal of Biological Chemistry* **235:** 1580-1585.

Howe, J.G. and Hershey, J.W.B. (1983) Initiation-factor and ribosome levels are coordinately controlled in *Escherichia coli* growing at different rates. *Journal of Biological Chemistry* **258:** 1954-1959.

Hu, W.S., Wang, R.Y.H., Shih, J.W.K. and Lo, S.C. (1993) Identification of a putative *infC-rpmI-rplT* operon flanked by long inverted repeats in *Mycoplasma fermentans* (incognitus strain). Gene **127:** 79-85.

Hui, A. and de Boer, H.A. (1987) Specialized ribosome system - preferential translation of a single messenger-RNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **84:** 4762-4766.

Iost, I. and Dreyfus, M. (1995) The stability of *Escherichia coli lacZ* messenger-RNA depends upon the simultaneity of its synthesis and translation. *EMBO Journal* **14:** 3252-3261.

Itoh, T. and Tomizawa, J-I. (1978) Initiation of replication of plasmid ColE1 DNA by RNA polymerase, Ribonuclease H, and DNA polymerase I. *Cold Spring Harbor Symposia on Quantitative Biology* **43:** 409-417.

Jacques, J.P. and Kolakofsky, D. (1991) Pseudo-templated transcription in prokaryotic and eukaryotic organisms. *Genes & Development* **5:** 707-713.

Jacques, N., and Dreyfus, M. (1990) Translation initiation in *Escherichia coli* - old and new questions. *Molecular Microbiology* **4:** 1063-1067.

Jain, C. and Belasco, J.G. (1995) RNase E autoregulates its synthesis by controlling the degradation rate of its own messenger RNA in *Escherichia coli* - unusual sensitivity of the *rne* transcript to RNase E activity. *Genes & Development* **9:** 84-96.

Kalapos, M.P., Cao, G.J., Kushner, S.R. and Sarkar, N. (1994) Identification of a 2nd poly(A) polymerase in *Escherichia coli*. *Biochemical and Biophysical Research Communications* **198:** 459-465.

Kang, P.J. and Craig, E.A. (1990) Identification & characterization of a new *Escherichia coli* gene that is a dosage-dependent suppressor of a *dnaK* deletion mutation. *Journal of Bacteriology* **172:** 2055-2064.

Kawamukai, M., Utsumi, R., Takeda, K., Higashi, A., Matsuda, H., Choi, Y. and Komano, T. (1991) Nucleotide sequence & characterization of the *sfs1* gene: *sfs1* is involved in CRP*-dependent *mal* gene expression in *Escherichia coli*. *Journal of Bacteriology* **173:** 2644-2648.

Kelly, K.O. and Deutscher, M.P. (1992) The presence of only one of five exoribonucleases is sufficient to support the growth of *Escherichia coli*. *Journal of Bacteriology* **174:** 6682-6684.

Keener, J. and Nomura, M. (1996) Regulation of Ribosome synthesis. In: Neidhardt, F.C., ed. *Escherichia coli and Salmonella: Cellular and Molecular Biology, 1*. Washington, DC: American Society for Microbiology. pp 1417-1431.

Kido, M., Yamanaka, K., Mitani, T., Niki, H., Ogura, T. and Hiraga, S. (1996) RNase E polypeptides lacking a carboxyl-terminal half suppress a *mukB* mutation in *Escherichia coli*. *Journal of Bacteriology* 178: 3917-3925.

Klotsky, R.A. and Schwartz, I. (1987) Measurement of *cat* expression from growth-rate-regulated promoters employing β-lactamase activity as an indicator of plasmid copy number. *Gene* **55:** 141-146.

Kohara, Y., Akiyama, K. and Isono, K. (1987) The physical map of the whole *Escherichia coli* chromosome - application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50:** 495-508.

Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes *Gene* **234:** 187-208.

Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of The National Academy of Sciences of The United States of America* **82:** 488-492.

Kushner, S.R. (1996) mRNA decay. In: Neidhardt, F.C., ed. *Escherichia coli and Salmonella: Cellular and Molecular Biology, 1*. Washington, DC: American Society for Microbiology. pp 849-860.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685.

Lateana, A., Pon, C.L. and Gualerzi, C.O. (1993) Translation of messenger-RNAs with degenerate initiation triplet AUU displays high initiation factor-II dependence and is subject to initiation factor-III repression. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 4161-4165.

Lin-Chao, S. and Bremer, H. (1986) Effect of the bacterial-growth rate on replication control of plasmid pBR322 in *Escherichia coli*. *Molecular & General Genetics* **203:** 143-149.

Lin-Chao, S., Chen, W-T. and Wong, T-T. (1992) High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Molecular Microbiology* **6:** 3385-3393.

Lin-Chao, S. and Cohen, S. N. (1991) The rate of processing and degradation of antisense RNA I regulates the replication of ColE1-type plasmids *in vivo*. *Cell* **65:** 1233-1242.

Littauer, U.Z. and Soreq, H. (1982) Polynucleotide phosphorylase. In: Boyer, P.D., ed. *The Enzymes*. New York: Academic Press. pp. 517-553.

Liu, C.G., Heath, L.S. and Turnbough, C.L. (1994) Regulation of *pyrBI* operon expression in *Escherichia coli* by UTP-sensitive reiterative RNA-synthesis during transcriptional initiation. *Genes & Development* **8:** 2904-2912

Liu, J.D and Parkinson, J.S. (1989) Genetics and sequence analysis of the *pcnB* locus, an *Escherichia coli* gene involved in plasmid copy number control. *Journal of Bacteriology* **171:** 1254-1261.

Liveris, D., Schwartz, J.J., Geertman, R. and Schwartz, I. (1993) Molecular-cloning and sequencing of *infC*, the gene encoding translation initiation factor-IF3, from four enterobacterial species. *FEMS Microbiology Letters* **112:** 211-216.

Lopez, P.J., Iost, I. and Dreyfus, M. (1994) The use of a transfer-RNA as a transcriptional reporter - the T7 late promoter is extremely efficient in *Escherichia coli* but its transcripts are poorly expressed. *Nucleic Acids Research* **22:** 1186-1193.

Lopez, P.J., Marchand, I., Yarchuk, O. and Dreyfus, M. (1998) Translation inhibitors stabilize *Escherichia coli* mRNAs independently of ribosome protection. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 6067-6072.

Lopilato J., Bortner, S. and Beckwith, J. (1986) Mutations in a new chromosomal gene of *Escherichia coli* K-12, *pcnB*, reduce plasmid copy number of pBR322 and its derivatives. *Molecular General Genetics* **205:** 285-290.

Maaloe, O. (1979) Regulation of the protein-synthesizing machinery - ribosomes, tRNA, factors and so on. In: Goldberger, R.F. ed. *Biological Regulation and Development, 1*. New York: Plenum Press. pp 487-542.

Mandel, M. and Higa, A. (1970) Calcium dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53:** 154.

March, J.B., Colloms, M.D., Hart-Davies, D., Oliver, I.R. and Masters, M. (1989) Cloning and characterisation of an *Escherichia coli* gene, *pcnB*, affecting plasmid copy number. *Molecular Microbiology* **3:** 903-910.

Marians, K.J. (1992) Prokaryotic DNA replication. *Annual Review of Biochemistry* **61:** 673-719.

Masai, H. and Arai, K-I. (1988) Initiation of lagging-strand synthesis for pBR322 plasmid DNA replication *in vitro* is dependent on primosomal protein i encoded by *dnaT*. *Journal of Biological Chemistry* **263:** 15016-15023.

Masters, M., Colloms, M.D., Oliver, I.R., He, L., MacNaughton, E.J. and Charters, Y. (1993) The *pcnB* gene of *Escherichia coli*, which is required for ColE1 copy number maintenance, is dispensable. *Journal of Bacteriology* **175:** 4405-4413.

Masters, M., Griffiths, C. and Charters, Y. (1992) Unpublished sequence data from the *dksA-pcnB* intergenic region submitted to the EMBL/GenBank databases in March 1992.

Masters, M., March, J.B., Oliver, I.R. and Collins, J.F. (1990) A possible role for the *pcnB* gene product of *Escherichia coli* in modulating RNA:RNA interactions. *Molecular General Genetics* **220:** 341-344.

Masukata, H. and Tomizawa, J-I. (1984) Effects of point mutations on formation and structure of the RNA primer for ColE1 DNA replication. *Cell* **36:** 513-522.

McCarthy, J.E.G. and Brimacombe, R. (1994) Prokaryotic translation - the interactive pathway leading to initiation. *Trends In Genetics* **10:** 402-407.

McCarthy, J.E.G. and Gualerzi, C. (1990) Translational control of prokaryotic gene-expression. *Trends In Genetics* **6:** 78-85.

McClure, W.R., Cech, C.L. and Johnston, D.E. (1978) A steady state assay for the RNA polymerase initiation reaction. *Journal of Biological Chemistry* **253:** 8941-8948.

Meinnel, T., Sacerdot, C., Graffe, M., Blanquet, S. and Springer, M. (1999) Discrimination by *Escherichia coli* initiation factor IF3 against initiation on non-canonical codons relies on complementarity rules. *Journal of Molecular Biology* **290:** 825-837.

Miller, J.H. (1972) *Experiments in Molecular Genetics*. New York: Cold Spring Harbor Laboratory Press.

Mohanty, B.K. and Kushner, S.R. (1999) Residual polyadenylation in poly(A) polymerase I (*pcnB*) mutants of *Escherichia coli* does not result from the activity encoded by the *f310* gene. *Molecular Microbiology* **34:** 1109-1119.

Mudd, E.A., Krisch, H.M. and Higgins, C.F. (1990) RNase E, an endoribonuclease, has a general role in the chemical decay of *Escherichia coli* mRNA: evidence that *rne* and *ams* are the same genetic locus. *Molecular Microbiology* **4:** 2127-2135.

Mulligan, M.E., Hawley, D.K., Entriken R. and McClure, W.R. (1984) *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. *Nucleic Acids Research* **12:** 789-800.

Nierman, W.C. and Chamberlin, M.J. (1979) Studies of RNA chain initiation by *Escherichia coli* RNA polymerase bound to T7 DNA. Direct analysis of the kinetics and extent of RNA chain initiation at T7 promoter A1. *Journal of Biological Chemistry* **254:** 7921-7926.

O'Hara, E.B., Chekanova, J.A., Ingle, C.A., Kushner, Z.R., Peters, E. and Kushner. S.R. (1995) Polyadenylylation helps regulate messenger-RNA decay in *Escherichia coli. Proceedings of the National Academy of Sciences of The United States of America* **92:** 1807-1811.

Petersen, C. (1992) Control of functional mRNA stability in bacteria: multiple mechanisms of nucleolytic and non-nucleolytic inactivation. *Molecular Microbiology* **6:** 277-282.

Polisky, B. (1988) ColE1 replication control circuitry - sense from antisense. *Cell* **55:** 929-932.

Pon, C.L., Brombach, M., Thamm, S. and Gualerzi, C.O. (1989) Cloning and characterization of a gene-cluster from *Bacillus stearothermophilus* comprising *infC, rpmI* and *rplT*. *Molecular & General Genetics* **218:** 355-357.

Portier, C. (1975) Quaternary structure of *Escherichia coli* polynucleotide phosphorylase: New evidence for a trimeric structure. *FEBS Letters* **50:** 79-81.

Powell, B.S., Court, D.L., Nakamura, Y., Rivas, M.P. and Turnbough, C.L. (1994) Rapid confirmation of single-copy lambda-prophage integration by PCR. *Nucleic Acids Research* **22:** 5765-5766.

Prentki, P. and Krisch, H.M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29:** 303-313.

Prentki, P. and Krisch, H.M. (1984) Transcription and translation termination by a selectable DNA fragment. *DNA* **3:** 199.

Pritchard, R.H., Barth, P.T. and Collins, J. (1969) Control of DNA synthesis in Bacteria. In: Meadow, P.M. and Pirt, S.J., ed. *19th Symposium of the Society for General Microbiology*. New York: Cambridge University Press. pp 263-297.

Py, B., Causton, H., Mudd, E.A. and Higgins, C.F. (1994) A protein complex mediating messenger-RNA degradation in *Escherichia coli*. *Molecular Microbiology* **14:** 717-729.

Py, B., Higgins, C.F., Krisch, H.M. and Carpousis, A.J. (1996) A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* **381:** 169-172.

Qi, F.X., Liu, C.G., Heath, L.S. and Turnbough, C.L. (1996) *In vitro* assay for reiterative transcription during transcriptional initiation by *Escherichia coli* RNA polymerase. *Methods in Enzymology* **273:** 71-85.

Qi, F.X. and Turnbough, C.L. (1995) Regulation of *codBA* operon expression in *Escherichia coli* by UTP-dependent reiterative transcription and UTP-sensitive transcriptional start site switching. *Journal of Molecular Biology* **254:** 552-565.

Raynal, L.C., Krisch, H.M. and Carpousis, A.J. (1998) The *Bacillus subtilis* nucleotidyltransferase is a tRNA CCA-adding enzyme. *Journal of Bacteriology* **180:** 6276-6282.

Reddy, P., Peterkofsky, A. and McKenney, K. (1985) Translational efficiency of the *Escherichia coli* adenylate-cyclase gene - mutating the UUG initiation codon to GUG or AUG results in increased gene-expression. *Proceedings of The National Academy of Sciences of The United States of America* **82:** 5656-5660.

Riley, M. (1993) Functions of the gene-products of *Escherichia coli*. *Microbiological Reviews* **57:** 862-952.

Robert-Lemeur M. and Portier C. (1992) *Escherichia coli* polynucleotide phosphorylase expression is autoregulated through an RNase III dependent mechanism. *EMBO Journal* **11:** 2633-2641.

Romero, A. and Garcia, P. (1991) Initiation of translation at AUC, AUA and AUU codons in *Escherichia coli*. *FEMS Microbiology Letters* **84:** 325-330.

Sacerdot, C., Chiaruttini, C., Engst, K., Graffe, M., Milet, M., Mathy, N., Dondon, J. and Springer, M. (1996) The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in *Escherichia coli*. *Molecular Microbiology* **21:** 331-346.

Sacerdot, C., Fayat, G., Dessen, P., Springer, M., Plumbridge, J.A., Grunbergmanago, M. and Blanquet, S. (1982) Sequence of a 1.26-kb DNA fragment containing the structural gene for *Escherichia coli* initiation factor-IF3 - presence of an AUU initiator codon. *EMBO Journal* **1:** 311-315.

Sambrook, J. Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (2nd Ed.). New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of The National Academy of Sciences of The United States of America* **74:** 5463-5467.

Seong, B.L. and RajBhandary, U.L. (1987) *Escherichia coli* formylmethionine transfer-RNA - mutations in GGG-CCC sequence conserved in anticodon stem of initiator transfer-RNAs affect initiation of protein-synthesis and conformation of anticodon loop. *Proceedings of the National Academy of Sciences of the United States of America* **84:** 334-338.

Shapkina, T.G., Dolan, M.A., Babin, P. and Wollenzien P. (2000) Initiation factor 3-induced structural changes in the 30 S ribosomal subunit and in complexes containing tRNA(f)(Met) and mRNA. Journal of Molecular Biology **299:** 615-628.

Sherratt, D.J. (1974) Bacterial Plasmids. *Cell* **3:** 189-195.

Simons, R.W., Houman, F. and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53:** 85-96.

Simons, R.W. and Kleckner, N. (1988) Biological regulation by antisense RNA in prokaryotes. *Annual Review in Genetics* **22:** 567-600.

Smith, R.W.P. (1995) The influence of the DnaA protein on transcription of the *ftsZ* and *dnaA* genes in *Escherichia coli.* Ph.D thesis, University of Edinburgh, Scotland.

Soberon, X., Covarrubias, L. and Bolivar, F. (1980) Construction and characterisation of new cloning vehicles. IV: deletion derivatives of pBR322 and pBR325. *Gene* **9:** 287-305.

Soderbom, F., Binnie, U., Masters, M. and Wagner, E.G.H. (1997) Regulation of plasmid R1 replication: PcnB and RNase E expedite the decay of the antisense RNA, CopA. *Molecular Microbiology* **26:** 493-504.

Sørensen, K.I., Baker, K.E., Kelln, R.A. and Neuhard, J. (1993) Nucleotide pool-sensitive selection of the transcriptional start site in vivo at the *Salmonella typhimurium pyrC* and *pyrD* promoters. *Journal of Bacteriology* **175:** 4137-4144.

Sprengart, M.L., Fatscher, H.P. and Fuchs, E. (1990) The initiation of translation in *Escherichia coli* - apparent base-pairing between the 16S rRNA and downstream sequences of the messenger-RNA. *Nucleic Acids Research* **18:** 1719-1723.

Sprengart, M.L. and Porter, A.G. (1997) Functional importance of RNA interactions in selection of translation initiation codons. *Molecular Microbiology* **24:** 19-28.

Steege, D.A. (2000) Emerging features of mRNA decay in bacteria. *RNA-A Publication of The RNA Society* **6:** 1079-1090.

Stormo, G.D., Schneider, T.D., Gold, L. and Ehrenfeucht, A. (1982) Use of the perceptron algorithm to distinguish translational initiation sites in *Escherichia coli*. *Nucleic Acids Research* **10:** 2997-3011.

Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage-T7 RNA-polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189:** 113-130.

Sussman, J.K., Simons, E.L. and Simons, R.W. (1996) *Escherichia coli* translation initiation factor 3 discriminates the initiation codon *in vivo*. *Molecular Microbiology* **21:** 347-360.

Tabor, S., and Richardson, C.C. (1985) Bacteriophage-T7 RNA-polymerase promoter system for controlled exclusive expression of specific genes. *Proceedings of The National Academy of Sciences of The United States of America* **82:** 1074-1078.

Takata, R., Mukai, T. and Hori, K. (1987) RNA processing by RNase III is involved in the synthesis of *Escherichia coli* polynucleotide phosphorylase. *Molecular & General Genetics* **209:** 28-32.

Talarico, T.L., Dev, I.K., Dallas, W.S., Ferone, R. and Ray, P.H. (1991) Purification and partial characterization of 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase and 7,8-dihydropteroate synthase from *Escherichia coli* MC4100. *Journal of Bacteriology* **173:** 7029-7032.

Talarico, T.L., Ray, P.H., Dev, I.K., Merrill, B.M. and Dallas, W.S. (1992) Cloning, sequence analysis, & overexpression of *Escherichia coli folK*, the gene coding for 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase. *Journal of Bacteriology* **174:** 5971-5977.

Tamm, J. and Polisky, B. (1985) Characterisation of the ColE1 primer-RNA I complex: analysis of a domain of ColE1 RNA I necessary for its interaction with primer RNA. *Proceedings of the National Academy of Sciences of the United States of America* **82:** 2257-2261.

Thanaraj, T.A. and Pandit, M.W. (1989) An additional ribosome-binding site on messenger-RNA of highly expressed genes and a bifunctional site on the colicin fragment of 16S ribosomal-RNA from *Escherichia coli* - important determinants of the efficiency of translation-initiation. *Nucleic Acids Research* **17:** 2973-2985.

Thomas, C.M. (1988) Recent studies on the control of plasmid replication. *Biochimica et Biophysica Acta* **949:** 253-263.

Tock, M.R., Walsh, A.P., Carroll, G. and McDowall, K.J. (2000) The CafA protein required for the 5 '-maturation of 16 S rRNA is a 5 '-end-dependent ribonuclease that has context-dependent broad sequence specificity. *Journal of Biological Chemistry* **275:** 8726-8732.

Tomizawa, J-I. (1984) Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell* **38:** 861-870.

Tomizawa, J-I. (1990) Control of ColE1 replication: Interaction of Rom protein with an unstable complex formed by RNA I and RNA II. *J. Mol. Biol.* **212:** 695-708.

Tomizawa, J-I. and Itoh, T. (1981) Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. *Proceedings of the National Academy of Sciences of the United States of America* **78:** 6096-6100.

Tu, A.H.T. and Turnbough, C.L. (1997) Regulation of *upp* expression in *Escherichia coli* by UTP-sensitive selection of transcriptional start sites coupled with UTP-dependent reiterative transcription. *Journal of Bacteriology* **179:** 6665-6673.

Twigg, A. J. and Sherratt, D. (1980) Trans-complementable copy-number mutants of plasmid ColE1. *Nature* **283:** 216-218.

Uhlin, B.E. and Nordstrom, K. (1977) R plasmid gene dosage effects in *Escherichia coli* K12: copy mutants of the R plasmid R1*drd19*. *Plasmid* **1:** 1-7.

van der Laken, K., Bakker-Steenveld, H. and van Knippenberg, P. (1979) Polyuridylic acid-dependent binding of fMet-tRNA to *Escherichia coli* ribosomes and incorporation of formylmethionine into polyphenylalanine. *FEBS Leters* **100:** 230-234.

Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., Krisch, H.M. and Carpousis, A.J. (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes & Development* **12:** 2770-2781.

Wagner, E.G.H. and Brantl, S. (1998) Kissing and RNA stability in antisense control of plasmid replication. *Trends in Biochemical Sciences* **23:** 451-454.

Wagner, E.G.H. and Simons, R.W. (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annual Review of Microbiology* **48:** 713-742.

Wagner, L.A., Gesteland, R.F., Dayhuff, T.J. and Weiss, R.B. (1994) An efficient shine-dalgarno sequence but not translation is necessary for *lacZ* messenger RNA stability in *Escherichia coli*. *Journal of Bacteriology* **176:** 1683-1688

Wang, R.F. and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene-expression in *Escherichia coli*. *Gene* **100:** 195-199.

Wickner, S.H. (1978) DNA replication proteins of *Escherichia coli*. *Annual Review of Biochemistry* **47:** 1163-1191.

Wright, J.J., Kumar A. and Hayward, R.S. (1992) Hypersymmetry in a transcriptional terminator of *Escherichia coli* confers increased efficiency as well as bidirectionality. *EMBO Journal* **11:** 1957-1964.

Wu, X.Q. and RajBhandary, U.L. (1997) Effect of the amino acid attached to *Escherichia coli* initiator tRNA on its affinity for the initiation factor IF2 and on the IF2 dependence of its binding to the ribosome. *Journal of Biological Chemistry* **272:** 1891-1895.

Xu, F.F. and Cohen, S.N. (1995) RNA degradation in *Escherichia coli* regulated by 3' adenylation and 5' phosphorylatian. Nature 374: 180-183.

Xu, F.F., Lin-Chao, S. and Cohen, S.N. (1993) The *Escherichia coli pcnB* gene promotes adenylylation of antisense RNA I of ColE1-type plasmids *in vivo* and degradation of RNA I decay intermediates. *Proceedings of The National Academy of Sciences of The United States of America* **90:** 6756-6760.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains - nucleotide-sequences of the M13mp18 and pUC19 vectors. *Gene* **33:** 103-119.

Young, R.A. and Davis, R.W. (1983) Efficient isolation of genes by using antibody probes. *Proceedings of The National Academy of Sciences of The United States of America-Biological Sciences* **80:** 1194-1198.

Yue, D.X., Maizels, N. and Weiner, A.M. (1996) CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. *RNA-a publication of The RNA Society* **2:** 895-908.

Zhang, J.R. and Deutscher, M.P. (1992) A uridine-rich sequence required for translation of prokaryotic mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 2605-2609.

Zilhao, R., Cairrao, F., Regnier, P. and Arraiano, C.M. (1996) PNPase modulates RNase II expression in *Escherichia coli*: Implications for mRNA decay and cell metabolism. *Molecular Microbiology* **20:** 1033-1042.

Zilhao, R., Regnier, P. and Arraiano, C.M. (1995) The role of endonucleases in the expression of ribonuclease II in *Escherichia coli*. *FEMS Microbiology Letters* **130:** 237-244.

Zubay, G. (1973) *In vitro* synthesis of proteins in microbial systems. *Annual Reviews of Genetics* **7:** 267-287.